PUPILLOMETRY AS A WINDOW ON THE ROLE OF MOTIONESE IN INFANTS' PROCESSING OF DYNAMIC ACTIVITY

by

JESSICA E. KOSIE

A DISSERTATION

Presented to the Department of Psychology
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

September 2019

DISSERTATION APPROVAL PAGE

Student: Jessica E. Kosie

Title: Pupillometry as a Window on the Role of Motionese in Infants' Processing of Dynamic Activity

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Psychology by:

Dr. Dare Baldwin	Chairperson
Dr. Caitlin Fausey	Core Member
Dr. Lou Moses	Core Member
Dr. Eric Pederson	Institutional Representative

and

Janet Woodruff-Borden	Vice Provost and Dean of the Graduate School

Original approval signatures are on file with the University of Oregon Graduate School.

Degree awarded September 2019

DISSERTATION ABSTRACT

Jessica E. Kosie

Doctor of Philosophy

Department of Psychology

September 2019

Title: Pupillometry as a Window on the Role of Motionese in Infants' Processing of Dynamic Activity

Over the first few years of life, infants acquire the ability to make sense of, predict, respond to, remember, and learn from a variety of everyday human actions. Finding segmental structure within unfolding activity – in particular, boundaries at which units of action begin and end – seems key to the acquisition of such action-processing fluency, and has important downstream implications for cognitive and linguistic development (e.g., Levine et al., 2018). However, action unfolds rapidly and is just as quickly gone. How do infants find structure in the complex, dynamic, fleeting action that they observe? Caregivers' infant-directed action demonstrations might serve to help with this challenging task. In interactions with infants, caregivers modify their motion in a variety of ways that engage infants' overall attention (i.e., "motionese;" Brand, Baldwin, & Ashburn, 2002). It seems likely that these modifications additionally highlight and promote infants' processing of the internal structure of action.

This dissertation explores the influence of motionese on infants' online processing of action. We first created a corpus of infant- and adult-directed activity sequences. Next, we use a recently-developed, open source, inexpensive, infant-friendly methodology to measure infants' pupil dilation as they viewed a select subset of these videos. We found

that infants' pupil size (an indication of attention or cognitive engagement) increased in response to action boundaries, but only for motionese demonstrations. Thus, in addition to engaging overall attention, motionese likely serves to promote infants' processing of action's internal structure. These findings set the stage for future work targeting the source of this increased pupil dilation at boundary regions.

In sum, this work makes several important contributions to developmental science. First, we have created a large, open video corpus of caregiver-infant interactions. We have also validated a new methodology for addressing any number of novel questions about infants' processing of visual information as it unfolds over time. Finally, this work provides the first demonstration to date that motionese influences infants' on-line action processing, and in this way scaffolds their understanding of, and ability to learn from, dynamic, novel activity.

CURRICULUM VITAE

NAME OF AUTHOR: Jessica E. Kosie

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene OR
Community College of Allegheny County, Pittsburgh, PA

DEGREES AWARDED:

Doctor of Philosophy, Psychology, 2019, University of Oregon
Master of Science, Psychology, 2014, University of Oregon
Bachelor of Science, Psychology, 2011, University of Oregon
Associate of Science, General Social Sciences, 2008, Community College of Allegheny County

AREAS OF SPECIAL INTEREST:

Developmental Psychology
Cognitive Development
Quantitative Methods
Open Science

PROFESSIONAL EXPERIENCE:

Graduate Teaching Fellow, University of Oregon Department of Psychology, 2012-2019

Graduate Research Fellow, University of Oregon Biology, Takahasi Lab, Summer 2015

Graduate Research Fellow, University of Oregon Psychology, Baldwin Lab, Fall 2011

GRANTS, AWARDS, AND HONORS:

Dissertation Research Fellowship, College of Arts and Sciences, University of Oregon, 2018 to 2019

Summer Dissertation Research Award, Institute of Cognitive and Decision Sciences, University of Oregon, 2018

Student Travel Award, Society for the Improvement of Psychological Science, 2018

Marthe E. Smith Memorial Science Scholarship, College of Arts and Sciences, University of Oregon, 2017

Graduate Education Committee Travel Award, Department of Psychology, University of Oregon, 2017

Marthe E. Smith Memorial Science Scholarship, College of Arts and Sciences, University of Oregon, 2016

Travel Award, Women in Graduate Sciences, University of Oregon, 2016

Beverly Fagot Dissertation Fellowship, Department of Psychology, University of Oregon, 2016

Graduate Education Committee Travel Award, Department of Psychology, University of Oregon, 2015

Graduate Education Committee Travel Award, Department of Psychology, University of Oregon, 2014

Alice C. Thompson Award for Outstanding Undergraduate Research, University of Oregon, 2010

PUBLICATIONS:

Baldwin, D. & Kosie, J. E. (under revision). How does the mind render streaming experience as events? *Topics in Cognitive Science.*

ManyBabies Consortium* (under revision). Quantifying sources of variability in infancy research using the infant-directed speech preference. *Advances in Methods and Practices in Psychological Science.*
*ROLE: data collection and data analysis

Kosie, J. E. & Baldwin, D. (2019). Attentional profiles linked to event segmentation are robust to missing information, *Cognitive Research: Principles and Implications 4*(1), 8.

Kosie, J. E. & Baldwin, D. (2019). Attention rapidly reorganizes to structure in a novel activity sequence, *Cognition, 182*, 31-44.

Kosie, J. E. & Baldwin, D. (2018). Tuning to the task at hand: Processing goals shape adults' attention to unfolding activity. *Proceedings of the 40th annual meeting of the Cognitive Science Society*.

Baldwin, D. & Kosie, J. E. (2018). Intersubjectivity and joint attention. *The International Encyclopedia of Anthropology*, 1-9. Wiley and Sons.

Kosie, J. E. & Baldwin, D. (2016). A twist on event processing: Reorganizing attention to cope with novelty in dynamic activity sequences. *Proceedings of the 37th annual meeting of the Cognitive Science Society*.

Kosie, J. E. & Baldwin, D. (2015). Flexibility is key in representational redescription. *The Newsletter of the Technical Committee on Cognitive and Developmental Systems, IEEE, 12*(2), 5-6.

ACKNOWLEDGMENTS

I would first like to thank the members of my dissertation committee, Professors Dare Baldwin, Caitlin Fausey, Lou Moses, and Eric Pederson for their support and input in designing this dissertation work. I am especially grateful to Professor Baldwin for nearly ten years of mentorship and collaboration. I would not be where I am today if I hadn't enrolled in her Psycholinguistics class so many years ago. I have also been fortunate to have had the opportunity to work with and learn from Professor Fausey. I am immensely appreciative of her guidance and many discussions about science and about academia more generally.

I would also like to thank Professors Sanjay Srivastava and Michael C. Frank. Professor Srivastava first introduced me to the concept of Open Science, and I continue to learn from him how to do good science in an open and replicable manner. I am grateful to Professor Frank who, in addition to serving as a role model for doing Open Science in the field of Developmental Psychology, has given me the opportunity to be involved in large-scale collaborative research and to share my Open Science skills with others.

I am grateful to my colleagues and friends in the Acquiring Minds Lab, past and present. I am particularly grateful to Dr. Jenny Mendoza for years of friendship and for listening to me complain when "everything is crap." This work would not have been possible without my talented team of undergraduate research assistants who put in countless hours recruiting participants, coding and entering data, and assisting with experimental sessions. Thank you as well to Dr. Avinash Bala whose continued assistance and support made it possible for me to use the SIPR (Stimulus-Induced Pupil Response) system in my dissertation research.

Special thanks to my wonderful husband, Dr. Shahar Shirtz, who is also my best friend and my biggest supporter. From troubleshooting bugs in my code to cleaning the litter boxes, I don't know how I would have survived graduate school and dissertation writing without him. (אני אוהבת את עוגיפלצת שלי!) I also cannot forget to thank my emotional support team – my cats, Rock Haim, Ravyn Byrd, Riley Bella, and Remy Jones.

Finally, I am incredibly appreciative of the families in the Eugene community who gave their time to participate in this work. This research could not exist without their generosity.

This dissertation is dedicated to my dad, John A. Kosie, without whom none of this would have ever been possible.

TABLE OF CONTENTS

Chapter Page

I. INTRODUCTION 1

Action Processing in Infancy 2

"Motionese" might scaffold infants' detection of structure in action 6

Limitations to Previous Research 10

Pupillometry as a Promising Solution 10

Overview of the Proposed Dissertation 15

Does Pupillometry Reveal Infants' Previously-Documented Preference for Motionese over Adult-Directed Action? 15

Do Infants Spontaneously Display a PDR in Relation to Action Boundaries? 16

Does Motionese Enhance Infants' Detection of Action Boundaries within Continuous Activity? 17

Open Science 18

II. CONSTRUCTING A VIDEO CORPUS OF INFANT- AND ADULT-DIRECTED ACTION 20

Introduction 20

The Kosie & Baldwin Video Corpus 21

Method 23

Participants 23

Materials 25

Filming Setup 26

Luminance Considerations 29

Chapter Page

Procedure 30

Final Stimulus Decisions 34

Coding 35

Luminance 38

Location of Action Boundaries 40

Infants' Object-Interaction Task 43

Use of Motionese 44

Results 45

Are there Variations in Luminance Throughout the Videos? 46

Is there Agreement Regarding the Location of Action Boundaries? 53

Do Infants Exhibit a Strong Preference for any of the Toys used in the Videos? 56

How Familiar did Caregivers Believe these Objects were to their Infants? 61

To What Extent did the Videos Depict Motionese? 64

Discussion 66

III. USING PUPILLOMETRY TO ASSESS THE INFLUENCE OF MOTIONESE ON INFANTS' PROCESSING OF DYNAMIC ACTIVITY 71

Introduction 71

Method 74

Participants 74

Apparatus 76

Design 77

Procedure 81

Chapter Page

Inclusion Criteria 83

Coding 84

Infant Gaze 84

Infants' Object-Interaction Task 85

Data Acquisition 86

Results 91

Validity checks: Did the Data Behave as Expected? 92

Did Motionese Enhance Infants' Overall Attention to Action? 99

Did Infants Selectively Attend to Action Boundaries in Continuous Activity Sequences? 102

Did Motionese Enhance Infants' Response to Boundaries within Continuous Activity? 104

Did Luminance Predict PDR Above and Beyond Effects of Demonstration Type and Region? 108

Was Infant Age Predictive of Looking Time and PDR Patterns Above and Beyond Effects of Demonstration Type and Video Region? 110

Did Infants Interaction with Objects More when they had Previously Viewed them in Motionese Demonstrations? 110

Discussion 116

IV. GENERAL DISCUSSION 124

A Corpus of Infant- and Adult-Directed Action 124

Comparison of Infants' Interest in Motionese Versus Adult-Directed Action 125

Motionese Facilitated Infants' Ability to Find Structure in Unfolding Activity 129

Limitations 133

Chapter Page

Broader Implications 137

Future Directions 139

Methodological Questions Raised in this Research 140

The Influence of Motionese on Infants' Overall Attention to Unfolding Activity 141

Exploring Infants' Processing of Action Boundaries in the Absence of Motionese 143

Further Investigation of Infants' Boundary-Related PDR to Motionese Activity 145

Conclusion 147

REFERENCES CITED 148

LIST OF FIGURES

Figure Page

2.1. Order of tasks in the corpus creation study 30

2.2. Luminance values for each of the twelve stimulus videos..... 48

2.3. Frames from the low- and high-luminance regions of the Infant-Directed Oball™ Stacker video 49

2.4. Average luminance at pre-boundary, boundary, and post-boundary regions..... 53

2.5. Agreement between experts and naïve research participants regarding the precise location of pre-specified action boundaries..... 55

2.6. Agreement between experts and naïve research participants regarding the precise location of action boundaries (participants not provided information about activity occurring at boundaries)..... 56

2.7. Proportion of trials on which infants first looked to each object..... 58

2.8. Proportion of three-second "looking" phase during which infants looked to each object 59

2.9. Proportion of the twenty-second "interacting" phase during which infants were interested in each object 60

2.10. Proportion of times a given object was subjectively coded as being preferred by infants..... 61

2.11. Caregivers' response to the question of whether or not the infant had seen the object before coming in to the lab 62

2.12. Average caregiver response to the question "How likely is it that your infant came into the lab today knowing what to do with this object?"..... 63

2.13. Average ratings for infant- and adult-directed demonstrations across each of the eight dimensions of motionese 66

3.1. Experimental setup..... 77

3.2. Structure of the pupillometry experiment 80

3.3. Proportion of time spent looking to stimulus videos across blocks..... 94

Figure Page

3.4. Average z-scored, filtered pupil size across the course of the videos..... 95

3.5. Average proportion of time infants spent looking to the video across the six blocks and by demonstration type 101

3.6. Average z-scored, filtered pupil size to infant- and adult-directed versions of each video..... 102

3.7. Average z-scored, filtered pupil size to motionese and adult-directed action..... 105

3.8. Average z-scored, filtered pupil size in response to motionese across the six blocks..... 108

3.9. Z-scored and filtered pupil size and z-scored video luminance to pre-boundary, boundary, and post-boundary regions of motionese and adult-directed action..... 109

3.10. Proportion of trials on which infants first looked to each object..... 112

3.11. Proportion of three-second "looking" phase during which infant looked to each object..... 113

3.12. Proportion of the twenty-second "interacting" phase during which infants were interested in each object..... 115

3.13. Proportion of trials in which infants "preferred" each object..... 116

LIST OF TABLES

Table Page

2.1. Maternal education across caregivers in the corpus study 25

2.2. Photos and suggested actions for each of the novel objects used in creation of the video corpus..... 27

2.3. Description of the twelve videos chosen for presentation to infants in the pupillometry study 36

3.1. Maternal education across caregivers in the pupillometry study..... 76

3.2. Total number of hits, missed, false alarms, and correct rejections across the seven Pi/Matlab coded and hand-coded videos..... 98

CHAPTER I

INTRODUCTION

Human activity generates a motion stream that is both complex and rapidly unfolding. Making sense of this dynamically streaming sensory information is a challenging cognitive enterprise; actions must be discerned "on the fly" as information streams past. The ability to find structure within unfolding activity (i.e., where individual units of action begin and end) is a key skill that is linked to fluency across domains including learning (Bailey, Kurby, Giovanetti, & Zacks, 2013), memory (Sonne, Kingo, and Krøjgaard, 2016, 2017; Sargent et al., 2013; Flores, Bailey, Eisenberg, & Zacks, 2017), social understanding (Zalla, Labruyére, & Georgieff, 2013), and language acquisition (Levine, Buchsbaum, Hirsh-Pasek, & Golinkoff, 2018). Early in life, infants seem to have acquired the ability to find structure in at least some kinds of activity sequences (see Levine et al., 2018 for a review). Less is known about how infants rise to the challenge of finding this structure as they first encounter novel action and watch it rapidly unfold over time.

It is important to consider, however, that infants don't face this challenge alone. Understanding the role of caregivers in early experience provides insight into the mechanisms that underlie infants' acquisition of complex cognitive skills like action processing. For example, in interactions with infants, caregivers modify their behaviors in a variety of ways that engage infants' attention and facilitate learning (Brand, Baldwin, & Ashburn, 2002; Fernald, 1985; Csibra & Gergely, 2009). It seems likely that caregivers' modifications to infant-directed action (e.g., "motionese;" Brand et al., 2002) could serve

specifically to help infants find structure as action unfolds. As yet this hypothesis remains untested, because methods with which to measure infants' online processing of streaming visual information have not yet been implemented in the action domain. However, the recent development of a novel, open-source, inexpensive, and infant-friendly system for measuring infants' pupillary response to cognitive stimuli (the SIPR (Stimulus-Induced Pupil Response) system; Patent Pending; Bala, Keller, Whitchurch, Baldwin, & Takahashi, 2016) provides a methodology with which to explore infants' online processing of visual information. The goal of this dissertation is to use the SIPR system to explore the extent to which motionese influences infants' ability to find structure as action unfolds across time.

In what follows, we first summarize what is currently understood about infants' processing of dynamically unfolding activity. Next, we discuss a small literature describing assistance that caregivers might provide to scaffold infants' processing of human action. Finally, we describe a body of evidence indicating that pupillometry offers potential insight into infants' processing of dynamically unfolding activity..

Action processing in infancy

A growing body of literature suggests that action segmentation processes are operative early in life (see Levine et al., 2018 for a recent review). In particular, infants display sensitivity to boundaries in a variety of everyday intentional action sequences (Baldwin, Baird, Saylor, & Clark, 2001; Hespos & Saylor & Grossman, 2009; Saylor, Baldwin, Baird & LaBounty, 2007; Hespos, Grossman, & Saylor, 2010). For example, in seminal work on action segmentation in infancy, Baldwin and colleagues (2001)

familiarized 10- and 11-month-old infants to a video depicting an actor engaging in a series of everyday activities. At test, infants were shown the same videos with pauses at action boundaries (i.e., the initiation and completion of intentional action units – like the moment at which one grasps an object to pick it up) or at non-boundary junctures. Infants looked longer to test videos that depicted pauses at non-boundary junctures, suggesting that they readily detect structure in unfolding intentional action, parse human behavior with respect to this structure, and are surprised when this structure is violated.

Recently, Sonne, Kingo, and Krøjgaard (2016) demonstrated that older infants' memory is influenced by the presence or absence of action boundaries, extending findings from studies with adults (e.g., Swallow, Zacks, Abrams, 2009; Radvansky & Zacks, 2017; Gold, Flores, & Zacks, 2017). In their research, one group of 16- to 20-month-old infants were shown action sequences with occlusions at boundary junctures while another group of infants saw action sequences with occlusions at non-boundary junctures. At test two weeks later, infants who were presented with stimuli that featured occlusions at boundaries had weaker memory for the activity than infants who were presented with stimuli featuring occlusions at non-boundary junctures. In an extension of this work, Sonne, Kingo, and Krøjgaard (2017) additionally demonstrated that, at a delay of ten minutes after viewing, 21-month-old infants more accurately remembered specific objects presented at action boundaries than those presented at non-boundary junctures. Results such as those described here provide evidence that infants, like adults, selectively attend to boundaries within unfolding activity. An open question, however, entails just how infants begin to find structure in dynamic action.

Statistical learning is one mechanism that seems likely to facilitate infants' ability to find structure in action. It has been demonstrated that infants can use the statistical regularities of extended action sequences to guide action segmentation at multiple levels of structure (Baldwin, 2012; Roseberry, Richie, Hirsh-Pasek, Golinkoff, & Shipley, 2011; Stahl, Roseberry, Hirsh-Pasek, Romberg, & Golinkoff, 2014). For example, 7- to 9-month-old infants viewed videos of hand movements (Roseberry et al., 2011) or an animated agent performing action sequences (Stahl et al., 2014). As in previous work with adults (e.g., Baldwin, Andersson, Saffran, & Meyer, 2008), these exposure corpora viewed by infants contained four different three-unit action sequences that were grouped into triads by the statistical regularities with which they co-occurred. *Within* a triad, each set of hand movements or animated actions always appeared in the same order as a unit (i.e., they had a transitional probability of 1.0 – if one movement occurred, it was 100% likely that the next movement in the triad would follow). In contrast, for items that occurred *across* the boundary between two triads the transitional probability was 0.5 (i.e., these movements occurred in sequence only 50% of the time). After being exposed to the corpus of actions, infants were shown sequences that depicted either statistically likely triads ("units" with a transitional probability of 1.0 between actions) or "part-units" that spanned the boundary between two action sequences. Across both studies, infants looked longer at the "part-units" suggesting that they had used the transitional probabilities to chunk action sequences into higher-level units and were surprised when test sequences violated this structure. This evidence suggests that infants as young as 7- to 9-months-old readily discover statistical structure within novel activity sequences; these results are

consistent with similar research with adult participants (e.g., Baldwin et al., 2008; Hard, Meyer, & Baldwin, 2018).

Several lines of evidence suggest that once infants have learned the predictability structure of action, they use this knowledge to guide their processing of unfolding activity (e.g., Ambrosini et al., 2013; Kanakogi & Itakura, 2011; Monroy, Gerson, & Hunnius, 2017). To illustrate, Monroy and colleagues (2017) familiarized 8- to 11-month old infants with a video that contained both random action sequences as well as action sequences with underlying statistical regularities (similar to those in the work by Roseberry, Stahl, and colleagues described earlier; Roseberry et al., 2011; Stahl et al., 2014). They monitored infants' gaze on a subsequent re-viewing of these sequences. Infants displayed an anticipatory gaze to the next action only in sequences that held the inherent statistical regularities, indicating that they had learned the structure of the activity sequence and were using this knowledge to predict what would occur next and guide their processing of the activity.

In sum, infants seem to be sensitive to the internal structure of at least some kinds of everyday activity, and their enhanced memory for activity occurring at boundaries suggests that they preferentially process these regions. Statistical learning is one likely mechanism that enables infants to discover the structure of action over time. However, it is unclear how much or what kind of repeated exposure is necessary before the statistics of a novel activity sequence can be learned and used to guide subsequent processing. In infants' day-to-day experience, some contexts might serve to enhance these statistics, promoting infants' identification of attention-worthy regions of activity (i.e., action boundaries), and thereby supporting infants' rapid acquisition of action processing skill.

One particular context that might be especially influential in this regard occurs when caregivers specifically attempt to demonstrate novel activities to infants.

"Motionese" might scaffold infants' detection of structure in action.

When demonstrating novel action to infants, caregivers modify their behavior in a variety of systematic ways that seem well suited to promoting infants' processing of the dynamic activity stream. Recent research investigating this phenomenon provides initial confirmation that infants benefit from such "motionese" demonstrations. It remains unclear, however, whether motionese specifically scaffolds infants' detection of structure within action, although this seems highly plausible.

A first study documenting motionese found that, when demonstrating novel objects to 6- to 13-month-old infants, caregivers exhibited increased interactiveness, proximity to their infant interactive partner, enthusiasm, range of motion, repetition, and simplicity in their actions (Brand et al., 2002). These modifications capture infants' attention, in that infants prefer to watch action demonstrations in a motionese format over action characteristic of demonstrations directed toward adults (Brand & Shallcross, 2008). Toddlers are also more likely to imitate actions demonstrated using motionese (Baldwin, Myhr, & Brand, in preparation; Williamson & Brand, 2014), and use of motionese increases 8- to 10-month-old infants' subsequent object exploration, which can have downstream benefits for overall learning (Koterba & Iverson, 2009).

The motionese modifications just summarized parallel modifications in language directed to infants, commonly called "motherese" (Snow & Ferguson, 1977), and are likely part of a suite of infant-directed modifications jointly constituting a natural

pedagogy phenomenon that has received extensive investigation in the developmental literature (Sage & Baldwin, 2010; Csibra & Gergely, 2009). Benefits of motherese in speech include facilitating infants' attention (Fernald, 1985), with subsidiary benefits such as enhancing infants' processing of the acoustic and segmental properties of speech (Kuhl, 2004), and promoting structure detection within streams of fluent speech (Thiessen, Hill, & Saffran, 2005; Kemler-Nelson, Hirsh-Pasek, Jusczyk, & Cassidy, 1989).

As with action, infants are sensitive to the statistical structure of language (e.g., Saffran, Aslin, & Newport, 1996; Aslin, Saffran, & Newport, 1998), and motherese appears to enhance infants' processing of these regularities. Thiessen and colleagues (2005) exposed 7-month-old infants to a novel, continuous syllable sequence with intonation contours characteristic of either adult-directed or infant-directed speech. Within the sequence, the only cues to word boundaries were statistical regularities across syllables; other characteristics of motherese (such as the length of pauses) that might influence infants' ability to recognize word-level units were equated across infant- versus adult-directed speech versions. They found that infant-directed intonation facilitated infants' detection of word-level units via statistical learning. In particular, in a subsequent test phase, infants who had heard the infant-directed version were better able to discriminate "words" (statistically predictable syllable sequences they'd previously heard) from "part-words" (sequences that spanned "word" boundaries) than those exposed to the adult-directed version. They concluded that infant-directed speech supported infants' detection of statistical structure in linguistic input. Given that the infant-directed speech in this research provided no direct clues to statistical structure,

these findings suggest that infant-directed speech assisted statistical learning by eliciting generally enhanced processing of the speech stream, an example of what is sometimes termed "social gating" (e.g., Kuhl, 2004).

In related research, Kemler-Nelson and colleagues (1989) explored the extent to which infants might be sensitive to naturally occurring prosodic cues within infant-directed speech as a source of information about the segmental structure of the speech stream. They hypothesized that prosodic features of clause boundaries that are characteristic of motherese speech (e.g., pauses, rising intonation, etc) might help infants segment the speech stream into clause-level units. Half of the 7- to 9-month-old infants in their study heard adult-directed speech and half heard infant-directed speech. In all speech samples, one-second pauses had been inserted either at clause boundaries or at within-clause locations. If infants are sensitive to prosodic cues as a source of information about clausal units, they should prefer speech in which pauses correlate with boundaries between these clausal units. Indeed, in the infant-directed condition, infants exhibited a preference for speech that contained pauses at clause boundaries, whereas pause location did not elicit any systematic difference in looking time for infants in the adult-directed condition. These results suggest that correlations between prosodic features of motherese and clause boundaries facilitated infants' detection of units within the complex speech stream.

Given such findings regarding *motherese*, it seems highly plausible that *motionese* analogously promotes infants' detection of structure within activity. In fact, there is existing evidence that certain features of motionese could serve to specifically direct infants' attention to action boundaries. For example, during object demonstrations to their

7- to 12-month-old infant, mothers' infant-directed gaze is systematically aligned with boundary junctures (Brand, Hollenbeck, & Kominsky, 2013). Features of mothers' infant-directed speech during action demonstrations is often aligned with action boundaries as well. For example, the onset and offset of mothers' action-describing speech tends to be aligned with boundaries occurring at the initiation or completion of an action unit (Meyer, Hard, Brand, McGarvey, & Baldwin, 2013; Hirsh-Pasek & Golinkoff, 1996), and infants tend to group such packaged action into coherent "chunks" (Brand & Tapscott, 2007). At action boundaries, mothers also tend to speak with rising or falling intonation, perhaps signaling the completion of an action unit (Rohlfing, Fritsch, Wrede, & Jungmann, 2006). Features such as repetition (Brand, McGee, Kominsky, Briggs, Gruneisen, & Orbach, 2009) and turn taking (Brand, Shallcross, Sabatos, & Massie, 2007) in infant-directed demonstrations occur systematically with action boundaries, and may additionally serve to facilitate infants' attention to the segmental structure of unfolding activity.

All in all, current evidence strongly suggests that motionese may assist infants in detecting action boundaries within continuously flowing activity, which would facilitate learning in a whole host of ways. For example, infants' ability to find structure in activity has possible downstream benefits for their ability to make sense of the action occurring around them (Zacks, Tversky, & Iyer, 2001), remember what has occurred (Sonne et al., 2016, 2017; Swallow et al., 2009), and perform actions themselves (Bailey et al., 2013). Infants' skill at detecting action boundaries would also promote social understanding (Zalla et al., 2013) and language learning (Levine et al., 2018). As yet, however, the

possibility that motionese scaffolds infants' detection of boundaries within streaming activity has not been put to direct test. This was a primary aim of the current dissertation.

Limitations to previous research

A critical barrier has stymied investigation into the extent to which features of motionese might direct infants' attention to boundaries within dynamic action. In particular, methods used in prior research provided little or no information about infants' moment-to-moment processing as activity unfolds. Instead, existing techniques for investigating infants' action processing have been largely limited to first exposing infants to action sequences and then, at later test, measuring infants' recognition/discrimination with respect to the stimuli that they previously viewed (e.g., Woodward, 1998; Baldwin et al., 2001; Stahl et al., 2014). Current understanding of the ways in which motionese influences infants' attention to unfolding action has been similarly constrained. Although existing research has clarified that motionese is preferred by infants and benefits their subsequent imitation of action, it has not been clear precisely how motionese influences infants' processing of action. However, a relatively new technique – measuring ongoing involuntary changes in pupil diameter concomitant with cognitive engagement – offers a novel approach to exploring issues related to infants' processing of unfolding action. This technique thus offers a novel window on ways in which motionese may scaffold such processing.

Pupillometry as a promising solution

Pupil dilation response (hereafter PDR) occurs spontaneously with changes in luminance (Loewenfeld, 1993) as well as in response to a variety of cognitive stimuli (Goldwater, 1972; Sirois & Brisson, 2014; Laeng, Sirois, & Gredebäck, 2012). Among other things, changes in pupil dilation are thought to reflect the attentional demands imposed by a cognitive task (Beatty & Lucero-Wagoner, 2000; Goldinger & Papesh, 2012). For example, adults' PDR increases with math problem difficulty (Hess & Polt, 1964) and as the number of items in working memory increases (Kahneman & Beatty, 1966; Peavler, 1974; Unsworth & Robinson, 2015). Further, pupil diameter is thought to track the degree to the allocation of attentional resources (Granholm, Asarnow, Sarkin, and Dykes, 1996; Granholm, Morris, Sarkin, Asarnow, & Jeste, 1997). Granholm and colleagues (1996) demonstrated that pupil diameter increased with the number of digits to be recalled, but only until participants' memory capacity was reached. At this point (i.e., when the number of digits to be recalled was approximately equal to participants' memory capacity) pupil diameter reached asymptote and then decreased as participants were asked to recall more digits than they could attend to at one time. PDR has additionally been used to index intensity of processing (Just & Carpenter, 1993), degree of mental effort (Kahneman & Beatty, 1966), surprisal (Preuschoff, t Hart, Einhäuser, 2011), response and orienting to novel or significant stimuli (Sokolov, 1963; Nieuwenhuis, De Geus, & Aston-Jones, 2014), and predictability of a stimulus (Nassar, Rumsey, Wilson, Parikh, Heasly, & Gold, 2013). Findings such as these are regarded as strong confirmation of Kahneman's (1973) suggestion that, among other things, pupil diameter provides an online indication of the "intensity of attention" being allocated by an observer.

Observed changes in pupil size are thought to be driven by activation in the locus coeruleus (LC), a subcortical structure that is considered the "hub" of the noradrenergic system (Aston-Jones & Cohen, 2005; Sara, 2009). The LC responds to stress by increasing secretion of norepinephrine and is linked to syndromes such as clinical depression, panic disorder, and anxiety (Carter et al., 2010; Klimek et al., 1997). It additionally appears to be involved in consolidation of memory (Sterpenich et al., 2006; Eschenko & Sara, 2008) and selective attention (Foote & Morrison, 1987). The linkage between LC activity and pupil dilation has been well established by studies using single-cell recordings in monkeys (e.g., Rajkowski, Majczynski, Clayton, & Aston-Jones, 2004; Joshi, Li, Kalwani, & Gold, 2016). A substantial body of evidence suggests that this link is present in humans as well. For example, low arousal states such as drug-induced drowsiness are characterized by both low tonic LC activity and reduced baseline pupil diameter (Morad, Lemberg, Yofe, & Dagan, 2000; Hou, Freeman, Langley, Szabadi, & Bradshaw, 2005). Additionally, processes thought to reflect LC activity, such as task engagement versus disengagement, correlate with changes in adults' pupil diameter (Gilzenrat, Nieuwenhuis, & Jepma, 2010; Jepma & Nieuwenhuis, 2011; Murphy, Robertson, Balsters, & O'Connell, 2011).

Modes of LC activity and corresponding pupil dilation are linked to two distinct patterns of behavior. Phasic activity occurs in response to observers' orientation to task-relevant stimuli and has been categorized as an "exploitation" mode of processing, while tonic activity reflects an "exploration" mode and corresponds to more general monitoring of the environment (see Laeng et al., 2012 for a review). For example, in a tone discrimination task, Aston-Jones and Cohen (2005) demonstrated that the pupil dilated

and subsequently restricted in response to each discrimination (i.e., phasic dilation) while baseline (i.e., tonic) pupil diameter continuously increased with task difficulty and peaked at the point at which participants decided to abandon the task and restart at a lower level of difficulty. Similar patterns of phasic and tonic response have been observed in response to linguistic stimuli as well. For example, Schluroff (1983) exposed adult participants to sentences varying in their linguistic organization and observed a phasic PDR to word onset as well as a tonic PDR to sentence difficulty. Specifically, overall average pupil size (tonic) increased with sentence difficulty, but across all levels of ambiguity there still a brief increase and return to baseline (phasic) response at the onset of each word in the sentence. In sum, while phasic dilation occurs in response to local stimuli relevant to the observer, tonic dilation occurs in response to general levels of task difficulty or arousal, though both tonic and phasic patterns of dilation can be observed in response to different features of the same stimulus.

Because the pupillary response is automatic, pupillometry enables the investigation of cognitive responses in nonverbal populations (e.g., Weiskrantz, Cowey, & Le Mare, 1998; Weiskrantz, Cowey, & Barbur, 1999). Recently, there has been renewed interest in the value of pupillometry in infancy research, and its use in this field has increased (e.g., Jackson & Sirois, 2009; Sirois & Jackson, 2012; Gredebäck & Melinder, 2010; Hepach & Westermann, 2016). With regard to this dissertation in particular, the use of pupillometry with infants offers a promising methodology with which to investigate the effects of motionese on infants' processing of dynamic human action.

Moreover, recent work from our research lab indicates that adults display systematic pupil dilation in relation to the internal structure of action sequences. Specifically, in a seminal study, Tanaka and colleagues (in preparation) presented adults with a series of short clips of sport activities, each containing one coarse-level action boundary (e.g., when the athlete completed their primary goal, such as striking a tennis ball with a racket during a serve). As predicted, we observed systematic changes in pupil diameter in relation to action boundaries. Adults' PDR was analyzed with respect to the time at which the major action boundary occurred within the videos. PDR tended to systematically increase immediately prior to action boundaries, peak at or shortly after boundaries, and return toward baseline over an extended period thereafter. This pattern of response indicates that the PDR methodology offers a window on viewers' detection of segmental structure within dynamic activity as processing is underway.

It seems plausible to predict that infants' PDR would display similar systematic relation to segmental structure within continuous activity sequences. For one, as described earlier, infants have been shown to be sensitive to the internal structure of at least some kinds of continuous action sequences; they can track statistical regularities inherent in extended action sequences and use these regularities to guide action segmentation at multiple levels of structure (Baldwin, 2012; Stahl et al., 2014; Monroy et al., 2017). Additionally, infants from as early as 4 months of age display systematic PDRs indicative of sensitivity to perceptual and goal structure (Jackson & Sirois, 2009; Gredebäck & Melinder, 2010; Sirois & Jackson, 2012; Addyman, Rocha, & Mareschal, 2014). A recently developed, inexpensive, open-source, infant-friendly PDR methodology, SIPR (Bala et al., 2016) made it possible to immediately undertake

investigation into the extent to which PDR provides an index of infants' detection of structure as action sequences unfold across time. This methodology additionally enabled us to examine the influence of motionese on infants' processing of dynamic activity.

Overview of the proposed dissertation

The overarching goal of this dissertation was to shed light on mechanisms that facilitate infants' processing of dynamic human action. This research addressed three main questions: (1) To what extent does infants' previously-observed preference for "motionese" over adult-directed action replicate via pupillometry (as opposed to standard looking-time measures as utilized in prior research)? (2) To what degree do infants spontaneously display systematic pupil-dilation response to action boundaries within streaming activity? and (3) To what extent does motionese specifically scaffold infants' detection of action boundaries within continuous activity sequences? A novel pupillometry paradigm makes it possible to investigate these questions for the first time. In the current study, infants viewed videos of motionese and adult-directed action as their pupil size was monitored. A secondary goal of this dissertation, therefore, was to validate a novel, inexpensive, open-source, and infant-friendly methodology that researchers can use to explore nuanced changes in the manner in which infants distribute their attention as they process streaming activity.

Does pupillometry reveal infants' previously-documented preference for motionese over adult-directed action? Infants are known to prefer to attend to motionese over adult-directed action (Brand & Shallcross, 2008). Other forms of natural pedagogy, specifically infant-directed speech or "motherese", similarly increase infants' arousal

(Fernald, 1985; Werker & McLeod, 1989; Cooper & Aslin, 1990). In pupillometry research, increased arousal manifests in increases in tonic (or sustained) PDR (Kahneman & Beatty, 1966; Laeng et al., 2012). Such sustained increase in PDR to high-arousal social stimuli has been observed across a variety of infant and preschooler research studies (e.g., Hepach, Vaish, & Tomasello, 2012 & 2015; Martineau, Hernandez, Hiebel, Roché, Metzger, & Bonnet-Brilhault, 2011; Geangu, Hauf, Bhardwaj, & Bentz, 2011; Nuske, Vivanti, Hudry, & Dissanayake, 2014 ; Nuske, Vivanti, & Dissanayake, 2015). We therefore predicted that the previously documented preference for motionese would be reflected in an enhanced tonic PDR to motionese action sequences relative to that observed in relation to the adult-directed action sequences. As a direct replication of previous research, we additionally measured infants' looking time to motionese and adult-directed activity (i.e., how long infants looked at the videos). Again, we expected to replicate prior research, predicting that infants would look longer to motionese than adult-directed action sequences.

Do infants spontaneously display a PDR in relation to action boundaries? A substantial body of prior evidence supports the prediction that, even in the absence of caregiver scaffolding, infants selectively attend to action boundaries, at least in some kinds of simple, familiar activity sequences. Specifically, prior research documents that infants detect boundaries within unfolding action (Baldwin et al., 2001; Saylor et al., 2007; Hespos et al., 2009) and display enhanced memory for content encountered at boundary regions relative to content occurring midstream within action units (Sonne, et al., 2017). These findings parallel research in adults (e.g., Newtson, 1973; Zacks et al., 2001; Hard et al., 2011; Kurby & Zacks, 2011; Richmond, Gold, & Zacks, 2017). As

described earlier, Tanaka and colleagues (in preparation) recently demonstrated that action boundaries elicit a systematic PDR in adults. Two previous sets of findings led us to predict that infants would display an analogous phasic PDR to action boundaries within unfolding activity: (1) infants' PDR profiles have shown a range of similarities to those of adults' (Jackson & Sirois, 2009; Gredebäck & Melinder, 2010; Sirois & Jackson, 2012), and (2) classic behavioral looking-time techniques have demonstrated that infants are sensitive to action boundaries in at least some simple, everyday activity sequences. Moreover, a study by Jackson and Sirois (2009) provides incidental evidence highlighting the plausibility of this prediction. They measured pupil diameter as infants viewed a train repeatedly entering and exiting a tunnel; infants' pupil dilation profiles displayed clear signs of a systematic PDR to the juncture at which the train exited the tunnel (that could not be explained by a change in luminance alone), which seems likely to coincide with what adults would judge to be an action boundary. Although investigating infants' PDR to action boundaries was not the focus of their research, their results nevertheless seem to provide evidence that infants exhibit a PDR in response to action boundaries, thereby increasing our confidence in predicting that infants would display a systematic PDR to action boundaries in human activity streams.

Does motionese enhance infants' detection of action boundaries within continuous activity? Speech modifications that are characteristic of motherese have been shown to enhance infants' ability to extract structure from dynamic streams of auditory stimuli (Thiessen et al., 2005; Kemler-Nelson et al., 1989). In the domain of action as well, caregivers appear to modify their behavior in ways that highlight action boundaries (e.g., Brand et al., 2013; Meyer et al., 2013; Rohlfing et al., 2006; Brand et al., 2009;

Brand et al., 2007). We thus expected to find a similar facilitative effect of motionese on infants' processing of action. That is, we predicted that, while infants would display an enhanced PDR to action boundaries even in the non-motionese condition, there would be a synergistic effect in that an increase in infants' PDR to boundaries would be larger when actions were demonstrated via motionese relative to adult-directed action.

In sum, we predicted an overall tonic effect of motionese, such that tonic PDR would be larger for actions demonstrated using motionese over an adult-directed format, that phasic responses to action boundaries would emerge across both motionese and adult-directed demonstrations, but that the phasic response to action boundaries would be larger when actions were demonstrated using motionese.

Open science

The movement to make scientific research more transparent and replicable is often referred to as "Open Science." Open Science practices include preregistering studies, hypotheses, and analysis plans, collecting adequately powered samples, making data open (i.e., accessible in an online data repository), and making code open (i.e., writing code in programming languages like R, ensuring that the code is thoroughly commented, and posting it online alongside the data) (see Klein et al., 2018 for a review). These practices enable researchers to outline their intentions prior to running a study, reducing the frequency of (even unintentional) questionable research practices – for example, doing a large number of analyses and reporting only those that were significant or peeking at one's data and collecting additional participants until the results reach significance. Open Science practices also increase the likelihood that, should another

researcher want to replicate the current study, they can readily do so. Results are more likely to replicate if the study was adequately powered to begin with and if the researcher intending to replicate the study knows exactly how the data were analyzed (i.e., has access to well-annotated code; see Hardwicke et al., 2018).

This dissertation project and analysis plans were preregistered on the Open Science Framework (OSF) website using a template provided by the OSF. The sample size that we intended to recruit for manuscripts resulting from this dissertation was carefully chosen and expected to adequately power the analyses of interest. Data were analyzed and figures created in R. The resulting, well-annotated, R scripts and corresponding data will be made available on GitHub and the OSF. At the end of each experimental session, we requested caregivers' consent to share the resulting data. For the infants from whom we received caregivers' consent (the majority of participants), videos of stimuli and experimental sessions will be made available online via Databrary.

CHAPTER II

CONSTRUCTING A VIDEO CORPUS OF INFANT- AND ADULT-DIRECTED ACTION

Introduction

The overarching aim of this dissertation work was to use pupillometry to investigate the extent to which caregivers' modifications to infant-directed action (i.e., "motionese") influenced infants' processing of dynamic activity sequences. To briefly summarize our goals for this pupillometry work, they were to: (1) investigate the extent to which motionese influenced infants' overall attention to unfolding naturalistic activity, (2) explore infant's overall ability to find structure in action as it unfolded in time, and (3) explore the influence of motionese on infants' attention to the structure of unfolding activity. To this purpose, we first created a set of videos of infant- and adult-directed action amenable to exploring these questions of interest. Our requirements for this set of videos were that they: (1) contain a set of short clips of caregivers acting on novel objects, with each clip containing one major action boundary (the precise location of which varied across clips), (2) include infant-directed and adult-directed clips that featured the same actor performing the same action on the same object, (3) maximize the difference in degree of motionese used in infant- versus adult-directed demonstrations, while (4) depicting action that seemed natural, and (5) equating as much as possible for luminance differences across videos. While some videos of infant- and adult-directed action already existed (e.g., Brand et al., 2002), no set of existing videos met all of the

above-mentioned criteria. Therefore, we created our own corpus of infant- and adult-directed action from which to select videos for the pupillometry study.

In addition to enabling us to explore infants' processing of action and the role of motionese, the creation of this corpus offered the opportunity to investigate a number of questions about the nature of caregivers' modifications to infant-directed action; questions such as the extent to which an individual caregiver's use of motionese and motherese is correlated, or the degree to which motionese differs in relation to infants' familiarity with a given object. Although investigation of these subsidiary questions was outside the scope of the dissertation, creation of stimuli was undertaken with the aim of being able to use these stimuli to address such questions in future research. In what follows we will briefly review considerations for creating this corpus and present a set of results describing various characteristics of the resulting set of videos that were used in the pupillometry study.

The Kosie & Baldwin Video Corpus

To create the video corpus, we asked adult caregivers of 9- to 18-month-old infants to demonstrate a set of objects to their infant and then to an adult. In everyday interactions with infants, most adults fairly naturally launch in to the use of natural pedagogy (Csibra & Gergely, 2009) which includes both motherese (e.g., Fernald, 1985) and motionese (e.g., Brand et al., 2002). However, in the absence of an infant, it turns out to be considerably more challenging to use motherese in language and motionese in action. Thus, while we considered using a single actor enacting scripted activity sequences to create stimuli for this research, we ultimately opted to collect a more

naturalistic stimulus set from caregivers interacting with their own infants, as in previous work in the domain of motionese (e.g., Brand & Shallcross, 2008). In this way, we were better able to increase the likelihood that the infant-directed action viewed by infants was as similar as possible to the sort of infant-directed action they might regularly be exposed to (assuming fairly homogenous behavior across caregiver/infant dyads).

Our methods for eliciting infant- and adult-directed action largely paralleled those of Brand and colleagues (2002), but differed in two main ways. In particular, in the previous research, caregivers were asked to bring a highly familiar partner such as a spouse, close adult friend, or their own mother. In the current research, however, we opted to ask caregivers to demonstrate objects to their own infant and to a research assistant. In contrast to the original work, in which one of the goals was to directly compare the use of motionese with infants and highly familiar adults, the primary goal of our corpus was to create a set of videos that maximally differed in the use of motionese. It seemed possible that using a highly familiar adult partner would decrease the difference in motionese between infant- and adult-directed action. For example, in the domain of speech, infant-directedness has been found in adults' speech to lovers or close friends (Trainor, Austin, & Desjardins, 2000) as well as to infants; thus, a highly familiar interaction partner might elicit more motionese than an unfamiliar partner during adult-directed demonstrations. Secondly, our corpus differed from that used in the original motionese research in that caregivers demonstrated objects to *both* their own infant and an adult partner. In the original research (Brand et al., 2002), caregivers demonstrated objects *either* to their own infants or to a highly familiar adult. In the interest of providing a first-ever documentation of the existence of motionese as a phenomenon of caregiver-

infant interaction, one goal in that research was to avoid caregivers becoming aware that the focus of the research was to compare infant- versus adult-directed action. That is, were caregivers to realize that the focus of the research was to compare infant- and adult-directed action, they might modify their behavior in ways that wouldn't have occurred naturally. However, now that spontaneous use of motionese has been documented and characterized to at least some degree, it seemed non-problematic for the present study to request caregivers to engage in both infant- and adult-directed demonstrations. Thus, for the current study, we opted to have all mothers produce action for the infant partner with a standard set of toys, and then to subsequently produce action for the adult partner with the same set of toys. Because of this ordering decision, it is possible that motionese could be reduced during the second demonstration simply because objects had become more familiar to the demonstrator. However, a possible reduction of motionese due to the demonstration to adults occurring second was not of concern given that our goal was to maximize differences in motionese between infant- and adult-directed action. Of course, for additional research using this corpus, these potential influences on adult- versus and infant-directed action will need to be clearly acknowledged and considered in any interpretation of the results.

Method

Participants

Fifty-three infants ranging from 9 to 18 months (29 females; Mean = 403 days; SD = 82.2 days) and their caregivers participated in stimulus filming. Motionese has been observed in caregivers' actions toward infants as young as 6 months of age (e.g., Brand et

al., 2002), and this is also the approximate age at which it has been documented that infants prefer motionese over adult-directed action (e.g., Brand & Shallcross, 2008). However, children appear to continue responding to motionese until at least two or three years of age (e.g., Williamson & Brand, 2014). Additionally, infants in the 9 to 18 month age range are starting to exhibit "secondary intersubjectivity" (Carpenter, Nagell, & Tomasello, 1999; Baldwin & Kosie, 2018; Trevarthen, 1977; Trevarthen & Hubley, 1978; Rochat, Passos-Ferreira, & Salem, 2009; Bakeman & Adamson, 1984), that is they begin to attend to relationships between people and objects rather than simply attending to people or objects alone. Finally, by 9 to 18 months of age, infants have the motor skills necessary to manipulate objects, and they begin to play with and explore toys (e.g., Lockman & McHale, 1989; Baldwin, Markman, & Melartin, 1993; Kimmerle, Mick, & Michel, 1995). Given these findings, infants between 9 and 18 months are likely to be interested in and responsive to object-focused demonstrations and their caregivers are likely to produce motionese in their infant-directed demonstrations.

Families from the local Eugene, OR community were recruited to participate through the University of Oregon Psychology Department's developmental database. Race/ethnicity of caregivers and infants was representative of the general Eugene, OR community. To assess socioeconomic status (SES), each family provided information about maternal education, a proxy for SES that tends to be predictive of developmental outcomes (e.g., Gottfried, Gottfried, Bathurst, Guerin, & Parramore, 2003; Noble, McCandliss, & Farah, 2007; Liaw & Brooks-Gunn, 1994). Mothers in our sample generally reported high educational achievement, with 42% reporting some level of

graduate training (see Table 2.1 for detailed information). After participating, all families received their choice of either a t-shirt or a children's book as a thank you gift.

Table 2.1.

Highest level of maternal education across caregivers in the corpus creation study. We report both the number of caregivers having achieved each level of maternal education as well as the proportion of the sample that this number represents. One caregiver did not provide maternal education information and is therefore missing from this summary.

Maternal Education	**Number of caregivers**	**Proportion of sample**
High School	4	7.5%
Some College	6	11.3%
Associate's Degree	2	3.8%
Bachelor's Degree	18	34%
Master's Degree	14	26.4%
Doctoral Degree	8	15.1%

Materials

The next step in corpus creation was to find a set of objects for caregivers to demonstrate to infants and adults. Again, our goal in object selection was to maximize the likelihood that caregivers would use characteristics of motionese in demonstrations to their infants. In line with the majority of previous research on motionese (e.g., Brand et al., 2002; Koterba & Iverson, 2009; Williamson & Brand, 2013; Brand & Shallcross, 2008), we opted to use novel, rather than familiar, objects. This decision was additionally supported by evidence that caregivers may be less likely to engage in motionese when their infant interaction partner already knows what to do with an object (Fukuyama et al., 2005).

To find these novel objects, we explored a variety of websites (e.g., Amazon) and local stores (e.g., Target, Fred Meyer, Dollar Tree) to amass a collection of objects that seemed likely to be relatively novel and interesting to infants. We then presented this collection of objects to adult faculty, graduate students, and undergraduate honors student

members of the Baldwin Acquiring Minds Lab and the Moses Developing Mind Lab. These individuals provided feedback on the objects that they felt were most novel and most likely to elicit motionese demonstrations. As a result of this feedback, we identified ten novel objects for use in creation of this corpus. These objects included: a pair of orange and green "fuzzy" shapes, a green tube that could be pulled to extend, a purple massage roller, a set of three Oballs™ that could be flattened and stacked, an Oball™ table-top toy with a suction cup on the bottom, a Duplo™ helicopter with removable pilot, a red silicone oven mit, a bright pink slinky, a blue, green, and orange ball covered with suction cups, and a multicolored circle that could be twisted. For each object, we additionally identified three or four actions that could be done with the object. (See Table 2.2 for photos and suggested actions for each object.)

Filming setup

During object demonstrations, caregivers were seated behind a half-circle shaped table and in front of a black curtain. For demonstrations directed to infants, infants were seated in a highchair approximately 4 feet away, facing their caregiver. A tripod was placed behind the infant, with a camera just above the infant's head that was focused directly at the caregiver. When demonstrations were directed to adults, the research assistant interaction partner sat on either a low stool, their knees, or the ground below the camera (depending on the research assistant's height) to ensure that their eyes were at approximately the same level that the infants had been. There was a mark on the table that aligned with the center of the camera, and caregivers were asked to try to remain centered with respect to that mark.

Table 2.2. *Photos and suggested actions for novel objects used in creation of the video corpus of infant- and adult-directed action. Starred objects are those used in the final set of video stimuli.*

Object	**Suggested Actions**
 **Fuzzy Shapes**	1. take pieces apart 2. spin ball piece on finger 3. put pieces back together
 **Green Tube***	1. stretch out tube 2. make tube into a circle 3. talk into end of tube 4. smash tube back together
 **Massage Roller***	1. roll across table 2. roll on palm of hand 3. hold up and spin roller
 **Oball™ Stacker***	1. smash balls by pushing on solid side 2. stack smashed balls 3. knock stack over

Table 2.2. Continued

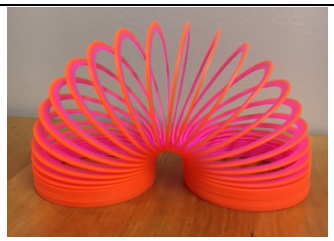
Object	Suggested Actions
 **Oball™ Table Toy**	1. stick to table 2. bend toy forward 3. spin ball
 **PD Copter**	1. spin blades of helicopter 2. fly helicopter straight up 3. fly helicopter around 4. take man out and put him back in
 **Red Mit**	1. place mit on hand 2. open and close hand inside mit 3. take mit off and demonstrate how to open/close fingers
 **Slinky***	1. pull up to stretch out 2. twist over to make “rainbow” shape 3. stretch out like an accordion 4. put on wrist

Table 2.2. Continued

Object	Suggested Actions
 **Sticky Ball***	1. stick to table and pop off 2. roll around on table 3. throw at table so it sticks
 **Twisty Glasses***	1. make into a figure 8 shape 2. place in front of eyes like glasses 3. open to make a circle, and put hand through

Luminance considerations

Because PDR is known to be impacted by luminance (e.g., Loewenfeld, 1993), we took a number of steps to minimize the luminance differences across infant- versus adult-directed videos. First, we used blackout window film on all of the windows in the study run room to ensure that the only light in the room came from the overhead light fixtures; thus lighting was equated across videos. All caregivers were seated in the same location, directly under one of the overhead light fixtures. Automatic exposure mode was disabled on the video camera to avoid automatic adjustments in exposure during the video recordings. We additionally took steps to control for visual differences between caregivers. All caregivers were asked to wear the same light blue t-shirt and, if their hair was long enough, to pull their hair back away from their face. Finally, caregivers had just

one object of interest on the table during any given demonstration. In these ways, we could best guarantee that overall brightness and general visual features of the videos would be as equivalent as possible across all demonstrations.

Procedure

As outlined in Figure 2.1, all infants participated in the ManyBabies study, watched as their caregivers demonstrated five pairs of objects (we refer to these as the "infant demonstration" tasks), and had an opportunity to interact with the objects themselves (the "infant interaction" tasks). Their caregivers additionally demonstrated the same five pairs of objects to an adult interaction partner (the "adult demonstration" tasks) and finally filled out a set of questionnaires while their infant played nearby.

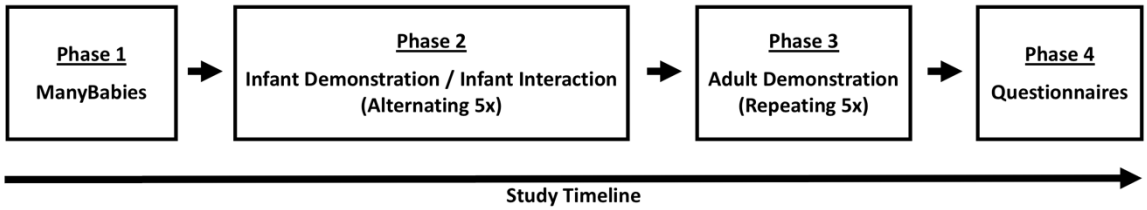


Figure 2.1. This figure depicts the order of tasks in the corpus creation study. Infants first participated in the ManyBabies study, then switched between viewing their caregivers demonstrating pairs of objects and interacting with the objects themselves. Next, caregivers demonstrated the same five pairs of objects to an adult partner before completing a set of questionnaires.

The ManyBabies study (ManyBabies Consortium, under revision) was a large-scale, collaborative replication of infant-directed speech preference that was unrelated to creation of our corpus. During this study, infants heard infant- and adult-directed speech. They controlled how long they heard the different speech streams by looking at a colorful centrally-presented digital checkerboard (i.e., if they looked away from the checkerboard

for more than two seconds, the speech stream stopped). Caregivers sat behind infants and listened to masking stimuli over noise-cancelling headphones while infants participated in the ManyBabies study. The masking stimuli consisted of music as well as infant- and adult-directed speech. While the ManyBabies study is generally unrelated to creation of the video corpus, it is important to note that before demonstrating objects to infants, caregivers did hear some infant-directed speech (as part of the masking stimuli). It is possible that hearing infant-directed speech could prime caregivers to engage in more motionese with their infant. However, because our goal was to maximize the difference between infant and adult-directed action, we were not concerned about a possible increase in motionese.

After completing the roughly six-minute ManyBabies task, caregivers and infants shifted position (described above) for the "infant demonstration" and "infant interaction" tasks. For the infant demonstration tasks, caregivers were given the following instructions: *"We're interested in how moms/dads demonstrate and talk about objects that infants haven't seen before. I have several objects that you'll show to [infant's name] one at a time. They're in pairs labeled 1 and 2. For each object, there's a card with some suggestions about things the object can do and that you might want to demonstrate."* If caregivers asked how long to play with the object, we told them approximately one minute, but we did not otherwise restrict the amount of time for which caregivers interacted with each object. After giving the instructions, an experimenter brought out the first set of objects and placed them on a low table next to the caregiver[1].

[1] Caregivers demonstrated objects in five sets of two objects each. For example, the first set of objects a caregiver might demonstrate would be Slinky then the Green Tube, the second the Twisty Glasses then the Massage Roller, and so on. Sets were ordered so that each object appeared in a particular place in the order (e.g., in set 1-5) an approximately equal number of times across participants. Orders were additionally

She then moved behind a curtain and remained there while caregivers demonstrated the objects to their infant.

Between each of the five infant demonstration tasks, infants participated in an "infant interaction" task. The purpose of this task was to explore infants' baseline levels of interest in the objects that were presented. In the infant interaction tasks, the experimenter placed the two objects that had just been demonstrated to the infant on a serving tray out of the infant's view and started a timer with a ticking clock sound. Then, the experimenter held up the tray so that the infant could see, but not reach, the objects and said *"Look what I have!"* She held the tray still for three seconds before saying *"Here you go!"* and placing the serving tray on top of the infant's highchair tray. Infants were then allowed to interact with the objects for another twenty seconds. The "infant demonstration" and "infant interaction" tasks were repeated in this way until the infant had viewed and interacted with all five sets of objects.

In the "adult demonstration" tasks, occurring next, caregivers were informed that they would now demonstrate the same actions on the same objects to an adult research assistant. The experimenter told them, *"Now [infant's name] is going to slide over here and play with me for a little while. I'm going to ask you to demonstrate the same objects and associated actions, but this time your interaction partner will be a research assistant who is already pretty familiar with these objects. Please show her these objects just as you would show any adult how to use them."* Once the infant was pulled aside and the

---

yoked across participants in an effort to control for the order of presentation within each object set. For example, another participant (yoked to the previously-described caregiver) would first demonstrate the Green Tube then the Slinky and, in their second demonstration, the Massage Roller then the Twisty Glasses. Thus, not only did objects appear in a particular place in the order an approximately equal number of times, but were also presented first and second about equally often.

research assistant was seated under the camera, the experimenter set the boxes of objects next to the caregiver. The boxes of objects were arranged in the same order in which they had been demonstrated to the infant. The experimenter then played, as quietly as possible, with the infant as the caregiver demonstrated all five object sets to the research assistant. While the experimenter and infant were playing, the infants' highchair was turned away from the demonstration that was occurring, except for a few instances in which the infant became very fussy when facing away from their caregiver. Research assistants were instructed to be friendly and responsive to the demonstrator, but to avoid engaging in lengthy conversations (e.g., they could smile and nod as felt natural, with only brief responses to any questions or comments from the demonstrator).

After the demonstration task, all caregivers completed three questionnaires. In the first questionnaire, they viewed images of the objects that they had played with and indicated if infants had seen each object before coming into the lab (replying yes, maybe, or no). If they replied "yes," or "maybe", they were asked to rate, on a scale of 1 to 5, how likely it was that infants would have come into that day's session knowing what to do with the object. Next, they completed a questionnaire asking for basic demographic information about their infant and their family. They also completed the Macarthur-Bates Communicative Development Inventory (Fenson et al., 1994; 2007) to assess the words and gestures that infants were comprehending and producing (this questionnaire was included for later use of this corpus; results are not reported in the current study). After completing the three questionnaires, caregivers were presented with a Databrary (Databrary, 2012) consent form and asked if they would be willing to allow us to share their videos with other researchers.

Final stimulus decisions

From the set of newly created digital videos described earlier (infant- and adult-directed action for each caregiver), we chose twelve clips (with six unique objects and six unique caregivers) to be used as stimuli for the pupillometry study. We had a number of criteria for choosing these twelve clips. First, only a handful of fathers participated; thus we chose to include only mothers in the final stimulus set. Our goal was to select videos in which infant- and adult-directed action were clearly distinct, while at the same time controlling for extraneous factors that might account for differences in infants' attention to adult and infant-directed demonstrations. Therefore, we opted to have the same actor depicted in both the infant- and adult-directed clips involving interaction with the same object. For example, mother A was featured in both the infant- and adult-directed action on the Sticky Ball while mother B was featured in both the infant- and adult-directed action on the Slinky. Because one of our goals was to optimize the chance that each infant in the pupillometry study would attend to multiple presentations of each clip, we opted to keep the clips fairly short (i.e., seven to twelve seconds in length). While we did allow length of the clips to vary across objects, the pair of infant- and adult-directed clips involving a given object were equated in length (see Table 2.3 for video descriptions including clip length). Finally, clips were selected and trimmed from full videos such that one major boundary (as defined by expert coders) occurred at approximately the same location within a pair of adult and infant-directed demonstrations. To summarize briefly, we chose a set of seven-to-twelve second clips of actors demonstrating the same actions on the same objects in infant- and adult-directed action; each video contained one major

action boundary, the location of which was matched across the infant- and adult-directed versions.

To choose the final set of video clips, all of the full-length demonstration videos were viewed, and a set of videos likely to best match the above criteria were selected. These were videos that contained one major action boundary, in which the difference between infant- and adult-directed action subjectively seemed most pronounced, and in which caregivers appeared to be performing similar actions in both the infant- and adult-directed demonstrations. From this set, we created a series of short clips and used subjective judgment to choose the twelve videos that we felt could be used to best address our questions of interest. These videos are available in a secure Databrary repository.

Coding

The twelve videos chosen for the pupillometry study were coded for a number of features. We first examined the potential of luminance-related influence on PDR effects. We next verified the expert coder's judgments regarding the location of action boundaries with two groups of naïve research participants. We additionally coded videos of the infants' spontaneous looking to, and play with, the objects during the infant interaction task to better understand infants' baseline interest in each of these items. Finally, a group of trained undergraduate research assistants coded all videos for the extent to which motionese characteristics were used in an effort to validate that the selected infant-directed videos were indeed representative of motionese and that the adult-directed videos were rated relatively lower in motionese features.

Table 2.3.

Description of the twelve videos chosen for presentation to infants in the pupillometry study. The table includes a still frame from each video, information about the interaction partner, the identity of the object, the total length of the video (in seconds and frames), the time at which the boundary occurred (in seconds and frames), and the activity that corresponded to the action boundary. Individuals making boundary judgments were told that the boundary occurred once the actor had finished the listed action.

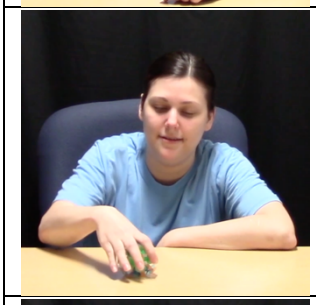
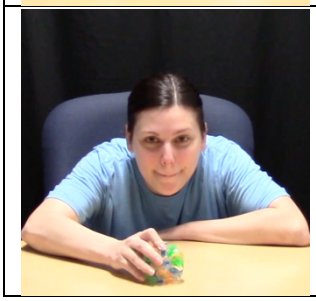
Still Image	**Partner Type**	**Object Name**	**Clip Length**	**Boundary Location**	**Boundary Description**
	Adult	Slinky	9 sec (270 frames)	1.93 sec (frame 58)	Putting hand through the slinky.
	Infant	Slinky	9 sec (270 frames)	2 sec (frame 60)	Putting hand through the slinky.
	Adult	Sticky Ball	7 sec (210 frames)	2.93 sec (frame 88)	Sticking ball to the table.
	Infant	Sticky Ball	7 sec (210 frames)	2.97 sec (frame 89)	Sticking ball to the table.

Table 2.3. Continued

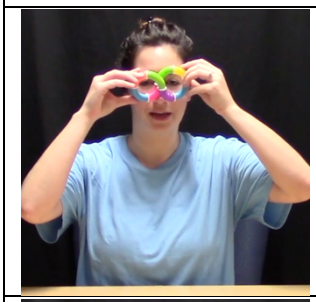
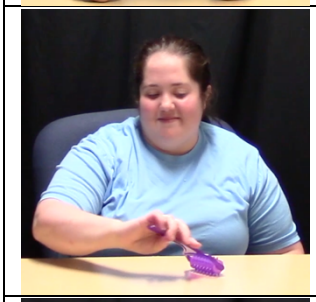
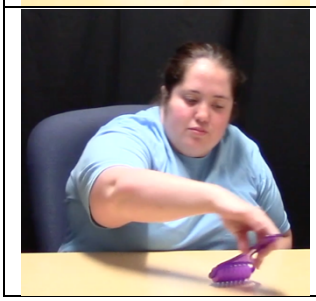
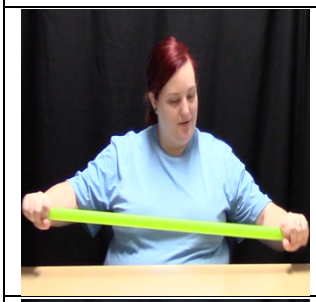
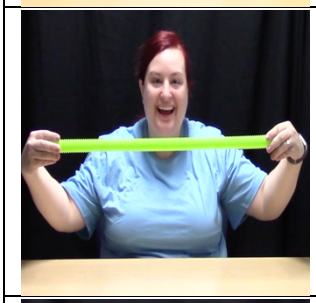
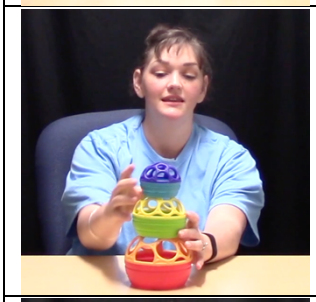
Still Image	Partner Type	Object Name	Clip Length	Boundary Location	Boundary Description
	Adult	Twisty Glasses	6 sec (180 frames)	3.6 sec (frame 108)	Bringing glasses to her face.
	Infant	Twisty Glasses	6 sec (180 frames)	3.93 sec (frame 118)	Bringing glasses to her face.
	Adult	Massage Roller	10 sec (300 frames)	3.8 sec (frame 114)	Rolling massager on the table.
	Infant	Massage Roller	10 sec (300 frames)	3.93 sec (frame 118)	Rolling massager on the table.

Table 2.3. Continued

Still Image	Partner Type	Object Name	Clip Length	Boundary Location	Boundary Description
	Adult	Green Tube	9 sec (270 frames)	5.8 sec (frame 174)	Stretching out Green Tube
	Infant	Green Tube	9 sec (270 frames)	4.06 sec (frame 122)	Stretching out Green Tube
	Adult	Oball™ Stacker	12 sec (360 frames)	9.77 sec (frame 293)	Stacking Oball™ toys.
	Infant	Oball™ Stacker	12 sec (360 frames)	9.6 sec (frame 288)	Stacking Oball™ toys.

Luminance. For each frame in each of our twelve videos, we used a weighted average of pixel values across the red, green, and blue (RGB) channels as an index of luminance. This method of calculating luminance is standard when working with video files (e.g., Poynton, 2003) and has been used to measure and control for the luminance of

a stimulus in prior infant pupillometry research (e.g., Jackson & Sirois, 2009; Hepach & Westermann, 2013; Geangu et al., 2011). First, we used a Matlab script (Matlab, 2019; script available in supplementary materials) to sum raw luminance values across each of the red, green, and blue channels for each frame in each of the twelve stimulus videos. Because these values were summed over the entirety of the 1920 x 1080 pixel-size frames (which thus each contain a total of 2,073,600 pixels), these values are very large. To put R, G, and B values for each frame into more standard units, we first divided them by the total number of pixels in the image. This gave us an average luminance for each of the R, G, and B channels, values for which fell in the standard range of 0-255. To make these values more interpretable, we additionally divided each value by 255 (also a standard procedure) so that the intensity of pixel values on the R, G, and B channels for each frame ranged from 0 to 1. Finally, before analysis, these values were corrected to reflect photometric luminance (e.g., luminance corrected for the perceived brightness by a human observer; Poynton, 2003). Specifically, the luminance of an RGB computer image can be measured via the intensity (as indexed by pixel value) of red, green, and blue channels. However, to use these values as a control for the luminance of a video, these values should be averaged and corrected for perceived brightness by a human (i.e., photometric luminance). We performed this correction on our luminance data, calculating a weighted average of the red, green, and blue channels with the following formula (Poynton, 2003):

$$Luminance = 0.2126 \times R + 0.7152 \times G + 0.0722 \times B$$

Thus, for each frame in each of the twelve stimulus videos, we had a single luminance measurement that reflected the weighted sum of the intensity (i.e., pixel value) of the red, green, and blue channels. These values ranged from 0 (a completely black image) to 1 (a completely white image).

Location of action boundaries. The location of action boundaries in each of the twelve videos was determined in three different ways. First, a single expert coder (Jessica Kosie) classified action boundaries in each of the stimulus videos. As described earlier, each video that was selected contained one major action boundary representing the completion of a coarse-level action unit. The expert coder viewed each of these videos and marked the frame number and seconds from the start of the video at which this boundary occurred. She then reviewed these judgments with a second expert coder (Dr. Dare Baldwin) who confirmed the location of these action boundaries. These expert coders have extensive experience engaging in research focused on action processing and, in particular, action segmentation. A still-frame from each video, description of action units, video length, and juncture at which the boundary occurred are available in Table 2.3.

These expert judgments were then further verified by two groups of naïve research participants. With these participants, we assessed level of agreement in two ways. We first described the activity occurring at the experimenter-defined boundary and asked participants to find the precise moment at which this boundary occurred in the stimulus videos. This served as validation of expert judgments of the precise moment at which the expert-defined boundaries occurred. We then assessed naïve participants' agreement with expert judgments of the location of action boundaries in the absence of

specific information about the activity content that occurring at each boundary. In this way, we addressed the question of whether naïve observers would nominate the same juncture in the activity sequence as a boundary.

To address the first question, participants (N = 63) were first provided with a brief description of how human observers process everyday actions and what it means to segment an action sequence (exact instructions available on the OSF page associated with this dissertation: http://osf.io/8mzhf). After this description, they were informed that they would watch a few sequences of action, focusing on one action unit in particular (for example, the actor rolling a massager on the table). After viewing the sequence once, they would be asked to view the sequence again and pause at the boundary marking the end of the specified action unit. They would then be asked to use the arrow keys to locate the exact moment they believed that the end of the action boundary had occurred. When participants indicated that they had found the precise moment at which the boundary occurred, a research assistant would record the location of that boundary in both frame numbers and milliseconds from the beginning of the video. We note here that while we instructed participants to find the "precise" moment that a boundary occurs, it is likely that observers' implicit processing of boundaries actually spans a broader time window surrounding the completion of an action unit (e.g., Kosie & Baldwin, 2019b). The intention behind asking participants to find precise moments within the activity was to encourage precision of these more explicit judgments of the end of action units that we could then use to probe implicit processing. Participants first completed two practice videos together with the research participant running the study, during which the research assistant provided feedback on the precise location of the action boundary. Then, for each

of the actual stimulus videos, participants were provided with instructions about exactly which boundary was the target of their segmentation judgment. For example, they might be told: *"In this video you'll see someone rolling a massager on the table and then rolling it on her hand. I'd like you to find the precise point in time at which she's completed the action of rolling the massager on the table."* After receiving these instructions, they viewed the video once in its entirety (without specifying the location of the boundary). Next, participants were reminded of the action unit of focus (e.g., *"Now, please find the precise point in time at which she's completed the action of rolling the massager on the table."*) and asked to play the video again, pausing at the action boundary, and using the arrow keys to find the moment at which they believed the target action had ended. Thus, the goal here was to validate the location of expert-specified action boundaries rather than to gather information about the juncture in the video at which naïve research participants believed a boundary had occurred.

In contrast, a second study with naïve research participants (N = 48) explored the extent to which participants agreed with experts' judgments about the location of the major action boundary, without receiving prior information about the content of activity occurring at the boundary location. As in the previous study, participants were first provided with a brief description of how human observers process everyday actions and what it means to segment an action sequence. They were then informed that they would watch each video once in real time. On a second viewing of the video, they would be asked to describe the activity content occurring at the action boundary. They then watched the video once more, pausing the video at the boundary, and using the arrow keys to locate the exact moment at which they believed the boundary occurred. The

research assistant would then record the moment at which the participant indicated a boundary had occurred. As in the previous study, participants completed three practice trials with guidance from the research assistant running the experiment. During these practice trials participants identified the location of the action boundary without prompting about its content and then received feedback that did include information about the precise location of the action boundary and the activity that was occurring at that region. However, no feedback regarding activity content or boundary location was provided after these practice trials. In contrast to the previous study, the goal of this study was to explore participants' judgments of the juncture in a video at which they believed a boundary had occurred rather than validating the timing of a pre-specified boundary.

Infants' object-interaction task. After each pair of objects had been demonstrated to infants, they were presented with that object pair. They first simply viewed the objects for three seconds on a serving tray just out of reach. The objects (still on the serving tray) were then placed on the infant's highchair tray and the infant interacted with the objects for twenty seconds. A set of trained research assistants coded these videos in two passes. First, two research assistants coded the first item the infant looked to when initially presented with the pair of objects. Specifically, research assistants paused the video of each object interaction immediately after the researcher said *"Look what I have!"* and specified the object to which the infant was looking. To assess inter-rater reliability, a subset (approximately 20%) of videos were double coded. Inter-rater reliability was high, with coders agreeing on the object of infants' first look on 92% of trials. In a second coding pass, using Datavyu (Datavyu Team, 2014), two trained research assistants coded the duration of infants' looking to each of the objects during the three-second looking-

alone phase, and the duration of infants' interest in each of the objects during the twenty-second interacting phase (coding instructions are available on the OSF webpage associated with this dissertation, http://osf.io/8mzhf). Coders were informed that interest in an object could include looking, touching, or manipulating an object, but to keep in mind that infants may be touching an object they aren't interested in or interacting with. Again, a subset of videos (approximately 20%) were double coded, and Cronbach's alpha was computed for the two sets of codes. Inter-rater reliability was again high for both the "looking-alone" and "interacting" phases, with Cronbach's alpha values of 0.96 and 0.97, respectively. During this coding pass, the two coders were additionally asked to make a subjective judgment about which toy was "preferred" throughout the time infants were looking and interacting with a given pair of objects. To assess inter-rater reliability on this measure, approximately 20% of videos were double-coded, and coders agreed on the identity of infants' "preferred" object on 97.5% of trials.

Use of motionese. The twelve stimulus videos were additionally coded for use of motionese using the dimensions of interest outlined by Brand and colleagues (2002). As in the original research, coders gave each demonstration a single rating (0-4) on each of eight global dimensions. Coders were provided with detailed instructions for coding each of the eight dimensions, including specific behaviors to look for (such as caregiver leaning forward and extending her arms toward the infant for the variable of *proximity to interaction partner*). The eight coded dimensions were: *repetitiveness* (0 = no repetitions, 4 = extremely repetitive); *rate* (0 = very slow, 4 = very fast); *punctuation* (0 = very fluid, 4 = very punctuated); *range of motion* (0 = very small, restricted movements, 4 = very broad, expansive movements); *proximity to partner* (0 = object never or almost never

leaves demonstrator's space, 4 = object always or almost always in participant's space), *simplification* (0 = complex combinations of many actions, 4 = small, simple units of action), *interactiveness* (0 = very low interaction, 4 = very high interaction), and *enthusiasm* (0 = very low enthusiasm, 4 = very high enthusiasm).

A set of trained undergraduate research assistants (N = 5) coded the twelve stimulus videos on all eight of the above-described dimensions. Coders were instructed to first watch the action demonstration in its entirety. They then watched the action demonstration again repeatedly with one of the eight dimensions in mind and coded only a single dimension at a time. All videos were coded with the sound off. It is important to note that, because videos were intentionally filmed without the demonstrator's interaction partner in the frame, coders could not see whether the interaction partner was an infant or an adult; thus coders were blind to whether a video exemplified infant versus adult-directed action. The full set of instructions given to coders are available on the OSF page associated with this dissertation (http://osf.io/8mzhf). Coders underwent training in this coding method prior to providing judgments for the twelve videos used in the pupillometry experiment. To "pass" this training, coders had to become reliable with expert judgments on a set of training videos. We considered RAs to be reliable when at least 90% of their judgments were within one point of expert judgments and when Cronbach's alpha for each dimension was greater than 0.64 (the minimum Cronbach's alpha reported by Brand et al., 2002).

Results

Our goals in the following analyses were to: (1) verify that the videos we chose for the pupillometry study did not systematically differ in ways that might influence our expected results, such as luminance differences, (2) collect baseline information about the salience of objects used in the videos, and (3) validate that actors' use of motionese varied along the expected dimensions in infant- versus adult-directed sequences. For analyses estimating linear mixed-effects models, we used the lme4 package (Bates, Mächler, Bolker, & Walker, 2015) in R (R Core Team, 2018) with type III sums of squares (set using the *afex* package; Singmann, Bolker, Westfall, & Aust, 2017). Significance for these models was assessed using the lmerTest package (Kuznetsova, Brockhoff, & Christiansen, 2015; Luke, 2017) with Satterthwaite's approximation for degrees of freedom. Following Barr, Levy, Scheepers, and Tily (2013), all models were fit with maximal random effects structure (intercepts and slopes) when possible; however, random slopes were removed when models failed to converge. The exact fixed and random effects structure that was used is specified for each model. For analyses requiring pairwise comparisons, we used the *lsmeans* package in R with a Bonferroni correction for multiple comparisons (Lenth, 2016). A standard alpha value of $p = .05$ was used to define statistical significance.

Are there variations in luminance throughout the videos?

In our first set of analyses, we were particularly focused on luminance features that might impact our pupillometry effects of interest. Therefore, we directly tested the extent to which luminance differed across videos of infant- and adult-directed action, across boundary, pre-boundary, and post-boundary regions, and whether any differences

at boundary regions might differ across infant- versus adult-directed activity sequences. To begin examining luminance differences across the stimulus videos, we plotted the luminance values of each frame separately by video and interaction partner (see Figure 2.2) and visually compared the extent to which they differed.

On visual inspection of the twelve graphs in Figure 2.2, luminance does not appear to differ dramatically across the demonstrations. As mentioned previously, the maximum possible luminance range with our corrected values was 0 (a completely black image) to 1 (a completely white image). In total across our twelve videos, the minimum luminance value was 0.259 and the maximum luminance value was 0.358, a difference of only 0.099. Within a given video, on average, the difference between the maximum and minimum luminance values was 0.027 (only about 3% of the maximum possible luminance difference), suggesting that variation in luminance within a given video was quite small. Therefore, it seems that luminance is unlikely to substantially influence participants' pupil diameter. In fact, Bala and colleagues (unpublished data) have demonstrated that even a 50% change in the luminance of a monitor, a much larger change than that which occurs anywhere in our stimulus videos, does not dramatically affect pupil dilation. However, we still opted to analyze luminance differences across infant- and adult-directed demonstrations and at boundary-related regions to control for any possible, though unlikely, luminance effects.

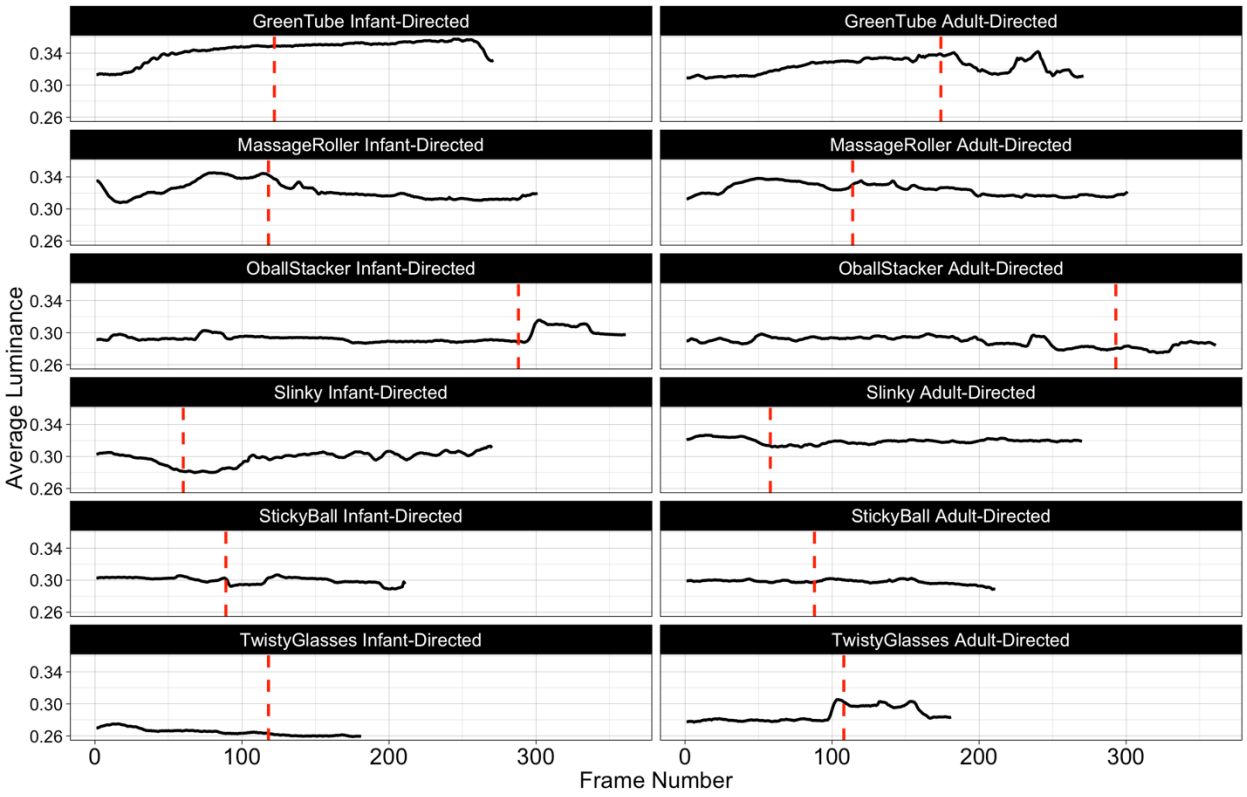


Figure 2.2. Luminance values for each of the twelve stimulus videos. Videos depicting the same object appear in the same row. All videos in the left-hand column are infant-directed and in the right-hand column are adult-directed. The location of the coarse-grained boundary in each video is indicated by the dashed red line.

When visual inspection of luminance patterns across a given video revealed pronounced changes in luminance (e.g., approximately frames 300 to 350 in the infant "Oball™ Stacker" video), we watched the video to better understand what kind of activity was occurring in that region and why these changes in luminance might have occurred. As illustrated in Figure 2.3, this inspection revealed that variations in luminance tended to correspond to movements of an actor's body, for example opening their arms. While pixel change (our method for measuring luminance) has been implemented as a way to track motion change in previous research (e.g., Loucks & Baldwin, 2009; Hard et al., 2011), the use of pixel change as an index of motion in the current stimuli appears to be tied to specific features of our videos. For example, the

actors are all wearing the same light-colored blue shirt, seated in front of a black background, and are lighter skinned. If any of these regions had been lighter or darker the observed luminance values would differ.

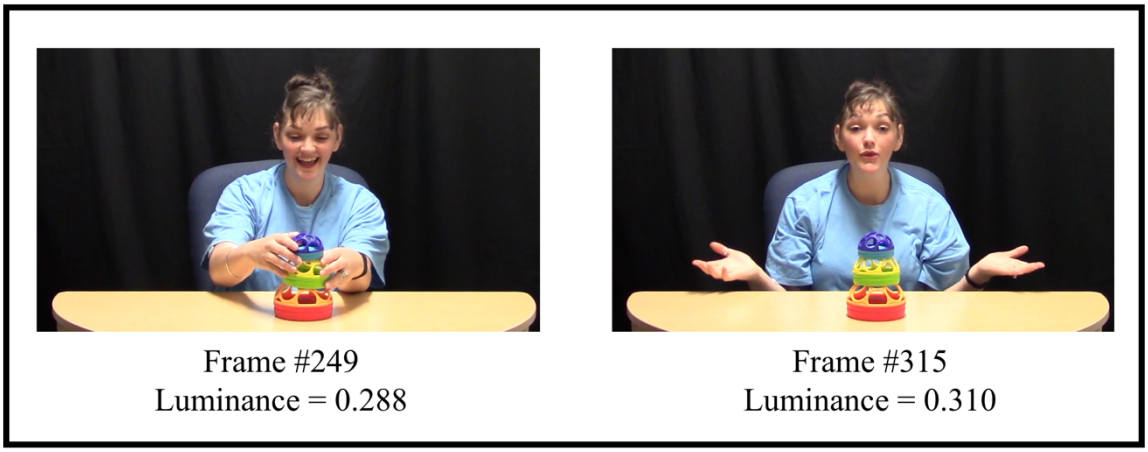


Figure 2.3. Frames from the low- and high-luminance regions of the Infant-Directed Oball™ Stacker video. As can be seen in the images, the increase in luminance appears to be due to the actor's hands taking up more of the black background on frame #315.

To examine the extent to which luminance might influence infants' pupil diameter as they viewed the stimulus videos, we first explored overall differences in luminance between infant- and adult-directed demonstrations. We opted to use corrected luminance values (i.e., corrected for photometric luminance following the steps outlined above) rather than z-scored values (used in later analyses) for this analysis as z-scoring could obscure mean-level differences across infant- and adult-directed demonstrations. We ran a linear mixed effects model predicting luminance from a fixed effect of interaction partner (infant vs. adult) and a random intercept for video. There were no significant differences in luminance across infant- ($M = 0.306$, $SD = 0.025$) and adult-directed action ($M = 0.308$, $SD = 0.018$), $\beta = 0.001$, $t(10) = 0.21$, $p = .84$, further supporting the

prediction that any observed overall effects of interaction partner in the pupillometry analysis are unlikely to be explained by luminance alone.

In our next set of analyses, we examined the potential influence of luminance on differences in pupil diameter across pre-boundary, boundary, and post-boundary regions. We also examined the extent to which these effects interacted with whether the interaction partner was an infant or an adult. Because we were not interested in mean-level differences between infant- and adult-directed action in these analyses, we opted to use z-scored luminance values. Z-scores were calculated separately for each of the twelve videos. Boundaries were defined using expert boundary judgments of the frame at which the one major action boundary in the activity sequence occurred (as mentioned previously, these judgments were verified by both a sample of trained undergraduate research assistants and a set of naïve research participants – the results of this verification process are described in the next section). From these judgments of the frame at which the boundary occurred, we defined boundary regions based on the same criteria used in our analyses of infants' pupil diameter (i.e., the analyses reported in Chapter III). Specifically, we defined pre-boundary, boundary, and post-boundary regions for each video. The *pre-boundary* region covered the one second of activity (or 30 frames) occurring prior to the action boundary. The *boundary* region began at the action boundary and extended for the next one second (30 frames), and the *post-boundary* region began one second post-boundary and continued one additional second, or 30 more frames. Because we were specifically interested in the region surrounding the boundary, the video frames included in these analyses were limited to those occurring in pre-boundary, boundary, and post-boundary regions. Video frames outside of these regions were

eliminated from the current analyses, though these frames were removed after z-scores were calculated (i.e., all frames for a given video, including frames outside of boundary-related regions, were included in our z-score calculations).

We first ran a linear mixed effects model predicting z-scored luminance from fixed effects of interaction partner (infant vs. adult), video region (pre-boundary, boundary, and post-boundary), and their interaction. We included a random intercept for video. Again, this model revealed no significant differences in luminance across infant- (M = 0.029, SD = 1.15) and adult-directed action (M = 0.051, SD = 1.08), $F(1, 10)$ = 0.002, p = .96. However, we did find a significant effect of video region, $F(2, 1{,}064)$ = 3.63, p = .03. To assess the simple effect of video region on z-scored luminance, we used the *lsmeans* package in R to compute pairwise contrasts with a Bonferroni correction for multiple comparisons (Lenth, 2016). The locus of this region effect seemed to be in the lower luminance values to boundary (M = -0.058, SD = 1.37) versus both pre-boundary (M = 0.090, SD = 0.99), β = 0.15, $t(1{,}064)$ = 2.35, p = .06, and post-boundary (M = 0.088, SD = 0.94) regions, β = -0.15, $t(1{,}064)$ = -2.32, p = .06. After Bonferroni correction, however, neither of these comparisons reached statistical significance at the p = .05 level. We additionally found no significant difference between pre-boundary and post-boundary regions, β = 0.002, $t(1{,}064)$ = 0.04, p > .99.

These results seemed best interpreted in the context of a significant interaction between partner and video region, $F(2, 1{,}064)$ = 8.90, p < .001. This interaction is depicted in Figure 2.4. To explore this interaction, we ran separate mixed-effects models for infant- and adult-directed action (with a fixed effect of video region and random video intercept). For adult-directed demonstrations, we found no difference in luminance

between pre-boundary (M = 0.118, SD = 1.05), boundary (M = 0.076, SD = 1.30), and post-boundary (M = -0.042, SD = 0.844) regions, $F(2, 532) = 1.83$, $p = .16$. However, as can be seen in Figure 2.4, luminance was (non-significantly) larger at boundary regions for adult-directed action, which is consistent with previous research using pixel values as an index of motion change at action boundaries (e.g., Hard et al., 2011). For infant-directed demonstrations we did find a significant effect of region, $F(2, 532) = 10.25$, $p < .001$, which was best described by a significant quadratic trend, $\beta = 0.28$, $t(532) = 4.20$, $p < .001$. Bonferroni-corrected pairwise contrasts revealed that luminance at the boundary region (M = -0.192, SD = 1.42) was significantly lower than luminance at the pre-boundary region (M = 0.062, SD = 0.93), $\beta = 0.25$, $t(532) = 2.78$, $p = .02$, and at the post-boundary region (M = 0.217, SD = 1.02), $\beta = -0.41$, $t(532) = -4.49$, $p < .001$. Luminance did not differ across pre- and post-boundary regions in the infant-directed demonstrations, $\beta = -0.155$, $t(532) = -1.70$, $p = .27$.

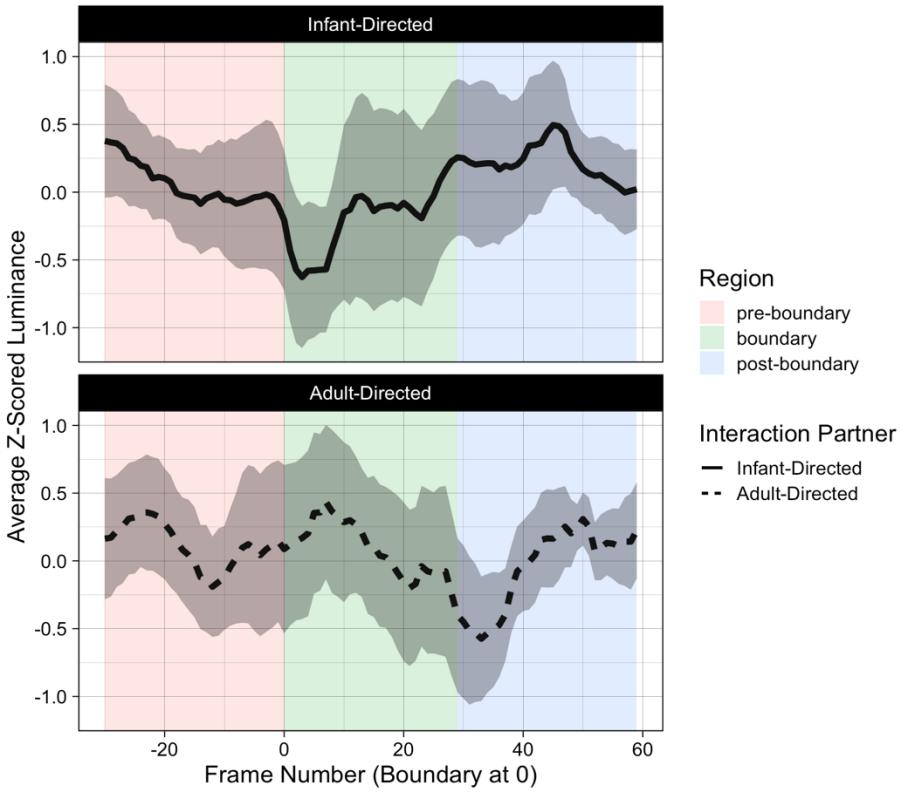


Figure 2.4. Average luminance at pre-boundary, boundary, and post-boundary regions. Here, the x-axis represents the region spanning the action boundary, with the action boundary occurring at time 0. The green shaded area represents the boundary region, or the 30 frames after the action boundary. The pink shaded area represents the pre-boundary region, 30 frames prior to the action boundary. The blue shaded area represents the post-boundary region, 30 frames after the end of the boundary region. Shading around lines indicates +/- 1 SE. This figure depicts all six adult-directed and all six infant-directed videos averaged together. Separate figures for each of the twelve videos are available in supplementary materials.

Is there agreement regarding the location of action boundaries?

Our next set of analyses focused on the location of the one major action boundary across the twelve videos in our stimulus set. While there are multiple ways to assess agreement about the location of action boundaries, we opted to explore the extent to which a naïve set of undergraduate research participants agreed on: (1) the timing of the point in the video at which an expert-defined action boundary occurred (e.g., the

boundary at which an actor had finished rolling a massager on the table), and (2) the point at which participants believed the action boundary occurred in the absence of information about what activity corresponds to the boundary location. For these analyses we defined the boundary region as the thirty frames (or one second) surrounding the point at which experts judged that the one major action boundary had occurred in each video; this definition of boundary region is consistent with related prior work (e.g., Zacks, Speer, Vettel, & Jacoby, 2006; Kurby & Zacks, 2011, 2018; Bailey et al., 2013). We then calculated the proportion of naïve undergraduate research participants who defined a boundary in that same thirty-frame (i.e., one second) bin.

In our first analysis, we examined the extent to which naïve research participants' (N = 63) judgments about the location of action boundaries corresponded to our expert judgments. These participants were provided with information about the activity occurring at the action boundary and were asked to find the exact moment at which that activity ended (e.g., *"Now, please find the precise point in time at which she's completed the action of rolling the massager on the table."*). As can be seen in Figure 2.5, agreement between experts and naïve participants was high, $r_{pb}(176) = 0.75$, $p < .001$, 95%CI[0.67, 0.81]. Participants were more likely to nominate a slide as a boundary if it fell into the same 1 second bins as the experts' boundary judgments than if it fell outside of that bin. When participants' judgments did not agree with expert judgments, they tended to fall after the expert judgment rather than before. A chi-square test confirmed that participants' judgments of the location of boundaries were more likely to be late (N = 418) than early (N = 217), $\chi^2(1) = 63.62$, $p < .001$.

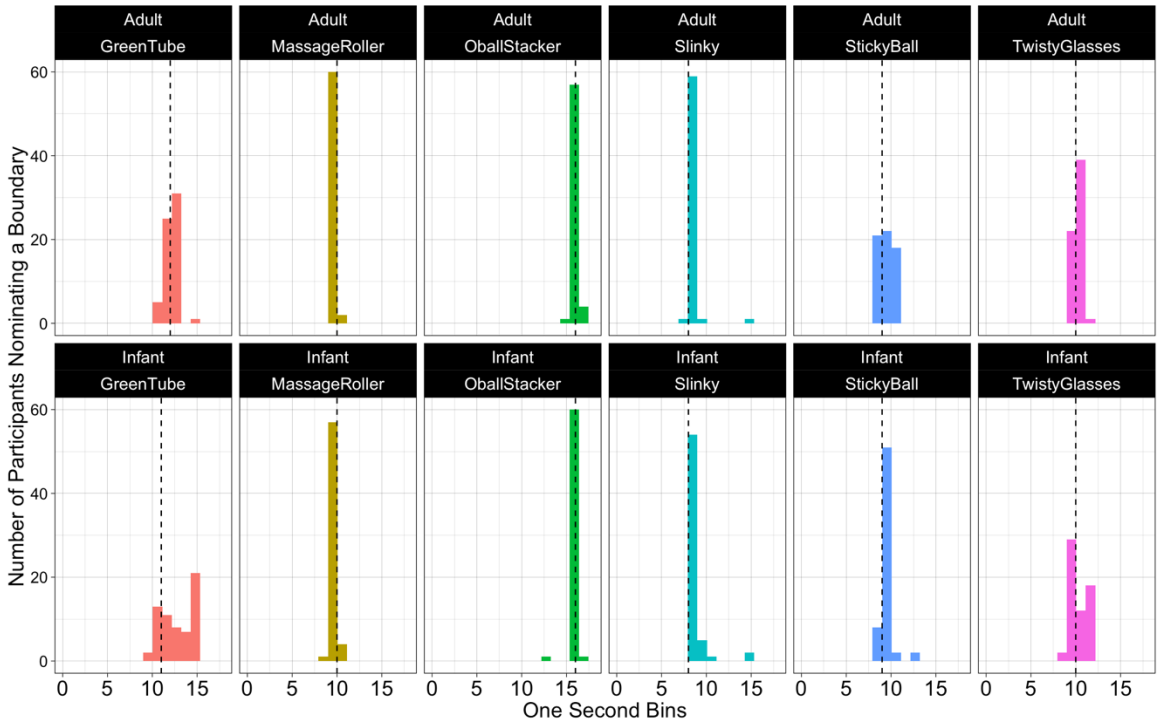


Figure 2.5. There was strong agreement between experts and naïve research participants regarding the precise location of pre-specified action boundaries in the twelve videos. Here, one-second bins are recorded on the x-axis. The y-axis represents the number of participants identifying a boundary in each bin. The color of the bars corresponds to the object being interacted with, and the location of the dashed line represents the expert boundary judgments.

Next, we again examined the extent to which naïve research participants' (N = 48) judgments about the location of action boundaries corresponded to our expert judgments. However, this time participants did not receive additional information about the activity content occurring at the boundary. They were simply asked to watch the video and decide where the one major action boundary occurred. Again, as can be seen in Figure 2.6, there was marked agreement between expert and participant boundary judgments, r_{pb} (176) = 0.68, $p < .001$, 95%CI[0.59, 0.75].

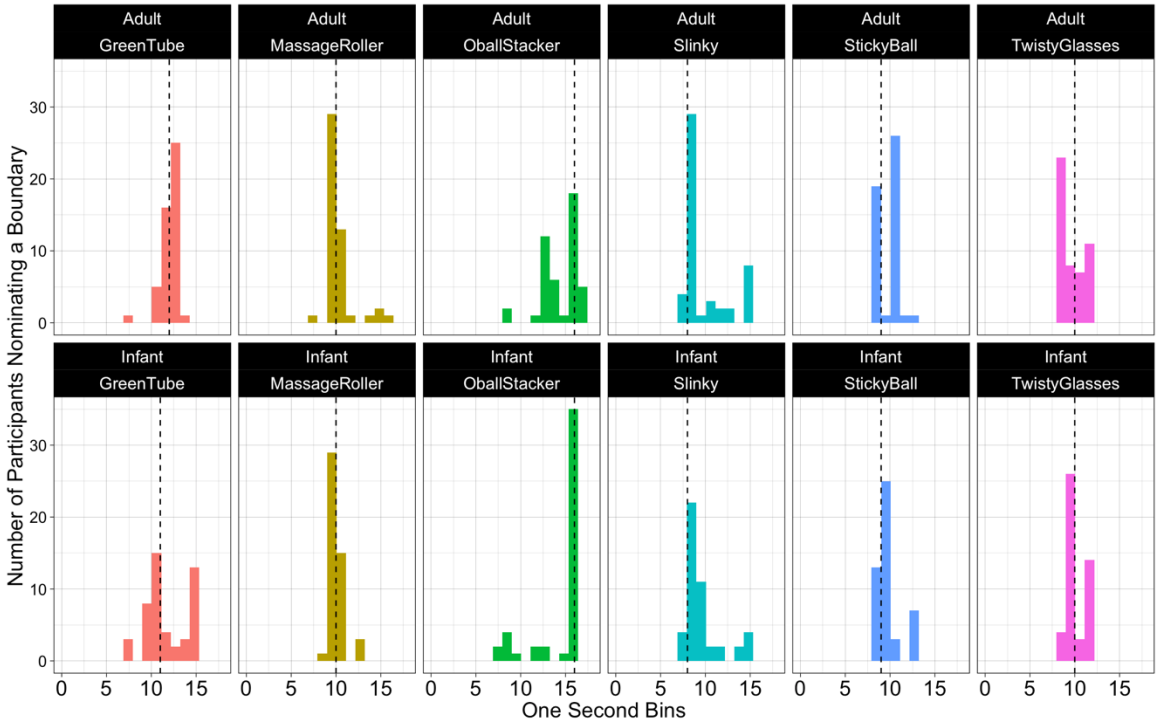


Figure 2.6. There was again strong agreement between experts and naïve research participants regarding the precise location of action boundaries in the twelve videos, even when participants were not provided information about the activity content occurring at boundaries. Here, one-second bins are recorded on the x-axis. The y-axis represents the number of participants identifying a boundary in each bin. The color of the bars corresponds to the object being interacted with, and the location of the dashed line represents the expert boundary judgments.

Even in the absence of information about the activity content occurring at action boundaries, participants were more likely to nominate a slide as a boundary if it fell into the same 1 second bins as the experts' boundary judgments. As with the previous analysis, when participants' judgments did not agree with expert judgments, they tended to late (N = 335) rather than early (N = 195), $\chi^2(1) = 3.98$, $p < .001$.

Do infants exhibit a strong preference for any of the toys used in the videos?

To get a sense of infants' baseline preference for the toys used in the pupillometry stimulus videos, we measured infants' interest in these objects in a variety of ways. As mentioned previously, during the "infant interaction" tasks, infants were shown the

objects in pairs. For a given pair of objects we coded the first object infants looked at after being presented with the pair, the length of time for which infants looked to each object in the first three seconds (the "looking-alone" phase), the length of time for which infants interacted with the object during the next twenty seconds (the "interacting" phase), and collected subjective judgments from coders regarding which object in each pair infants seemed to prefer. As with previous analyses, infants interacted with all ten objects used in stimulus filming. However, we report results from only the final set of six objects (i.e., those depicted in Table 2.3).

We first examined the proportion of times a given object was the target of infants' first look. This proportion for each object is depicted in Figure 2.7. A chi-square test revealed that there was some variability in the identity of objects that infants looked to first, $\chi^2(5) = 13$, $p = .02$. Specifically, the Oball™ Stacker was the most highly preferred object, with infants looking to it first on 65% of the trials in which it was presented. The least preferred toy, when defined as the first toy to which infants looked, was the Green Tube. Infants looked to this toy first on only 23% of the trials on which it was presented. As mentioned previously, the objects that were paired differed across infants (e.g., while one infant might get the Green Tube paired with the Slinky another infant might get the Green Tube paired with the Massage Roller). While it would be interesting to explore the proportion of time a particular object was preferred when paired with another specific object, our sample size is not large enough to facilitate this comparison. However, figures depicting the number of times an object was chosen when paired with each other object are included in supplementary materials.

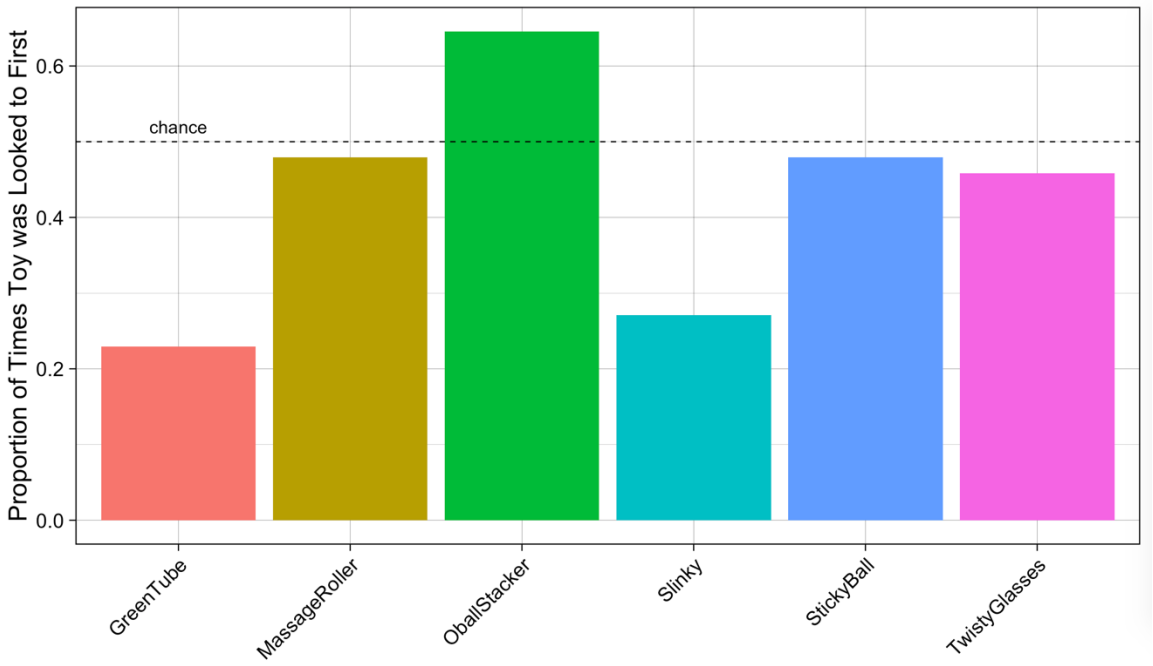


Figure 2.7. Proportion of trials (in which a given toy was presented) on which infants first looked to each object. The dashed line represents chance, which is .5 for any given object.

After infants had the opportunity to look at one of the objects, the tray with the given object pair was held just out of infants' reach for three seconds, and we coded the duration of their looking to each object during that time period. To examine overall differences in the proportion of time infants spent looking to toys, we ran a linear mixed-effects model predicting the number of seconds spent looking at a toy from the toy identity (a fixed effect) and random intercepts for subjects and object pair order (i.e., first, second, third, fourth, or fifth pair presented). We found that there were overall significant differences in the length of time for which infants looked to each object, $F(5, 211.62) = 5.06$, $p < .001$. These differences can be observed in Figure 2.8. Again, while infants did look at some objects for longer durations than others during the three-second "looking-alone" phase, there does not appear to clearly be one object that was overwhelmingly

preferred or ignored. As with the previous analysis, it would be interesting to explore the duration of looking to a given object when paired each other specific object. However, as mentioned previously, our sample size is not large enough to facilitate this comparison. Figures are again included in the supplementary material in which these comparisons can be visually examined.

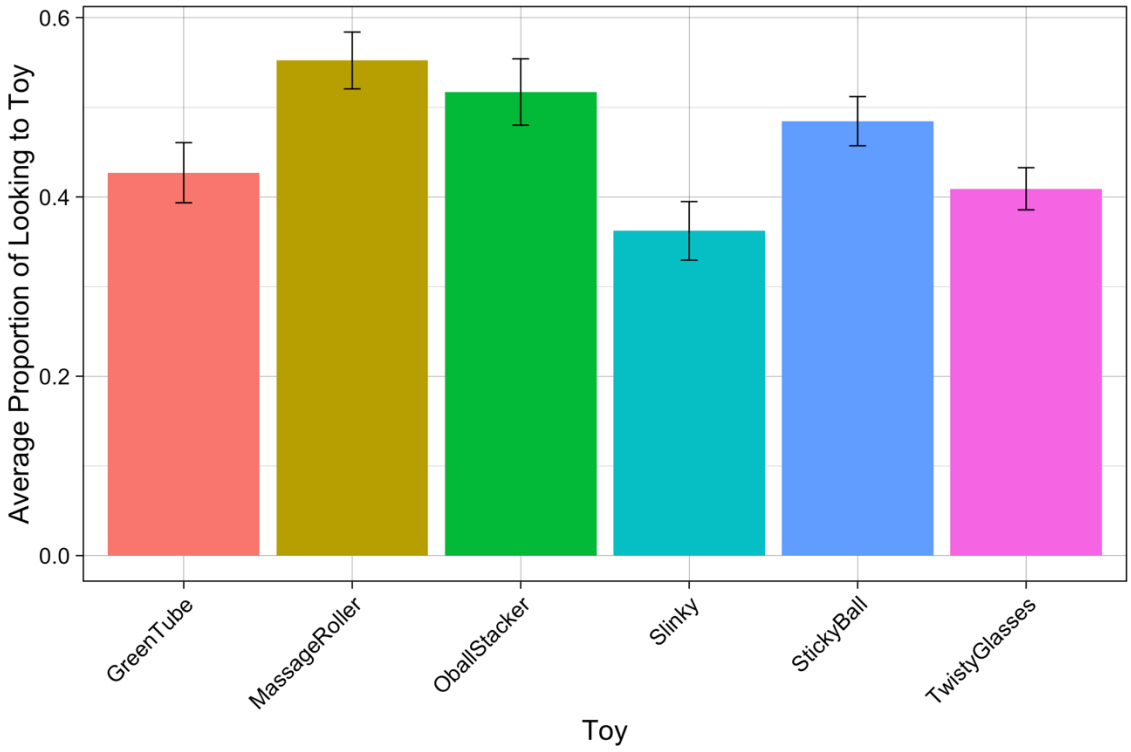


Figure 2.8. Proportion of three-second "looking" phase during which infants looked to each of the six objects included in the pupillometry stimuli. Error bars indicate +/- 1 SE.

After the three-second "looking-alone" phase, the serving tray containing the two objects was placed on the highchair tray in front of the infant and the infant was allowed to manipulate the objects for an additional twenty seconds. We refer to this as the twenty-second "interacting" phase. To test for differences in the length of time infants spent interacting with each object, we again ran a linear mixed-effects model predicting the number of seconds spent interacting with a toy from the toy identity (a fixed effect) and a

random intercept for subjects (including an additional random intercept for object set as we did in the previous analysis caused issues with model convergence, so it was omitted from this analysis). We again found significant differences in the duration of infant interactions with each object, $F(5, 260) = 6.34$, $p < .001$, as can be seen in Figure 2.9.

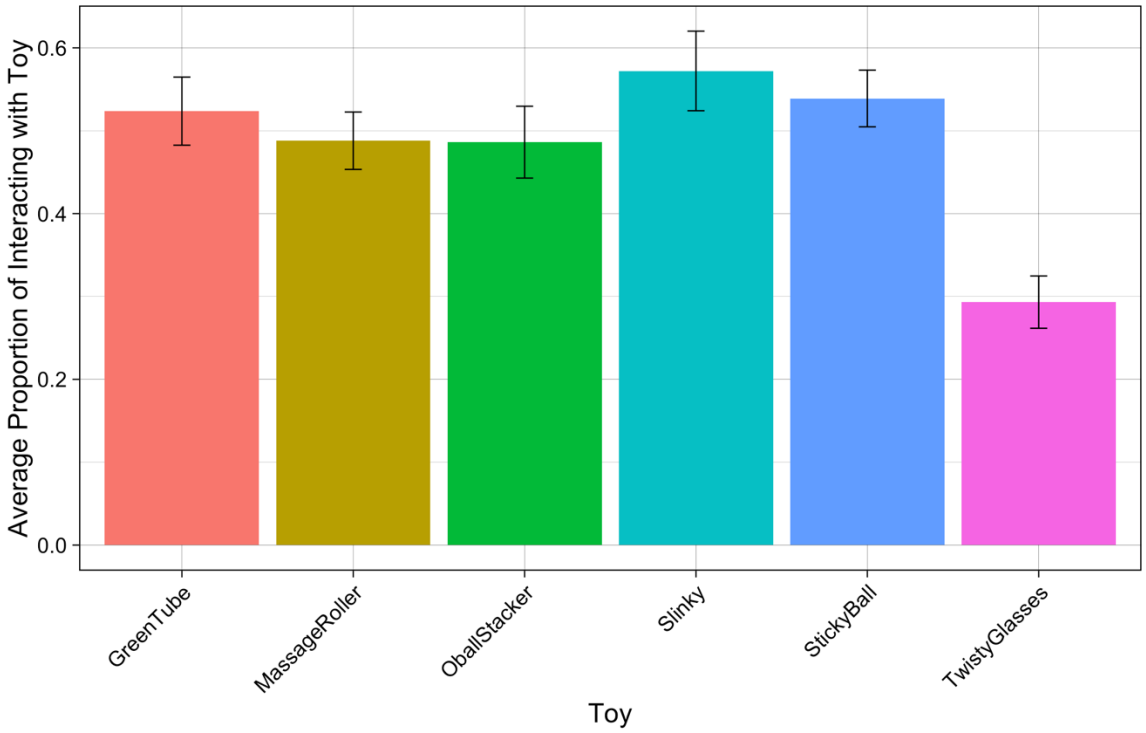


Figure 2.9. Proportion of twenty-second "interacting" phase during which infants were interested in each of the six objects included in the pupillometry stimuli. Error bars indicate +/- 1 SE.

While no single object stands out as having been overwhelmingly preferred, this analysis does suggest that that infants tended to be less interested in the Twisty Glasses, as they didn't spend much time interacting with this object. As with previous analyses, a figure is included in the supplementary materials where the duration of the "interacting" phase infants spent with each toy can compared for specific object pairs.

In our last analysis of the coded video data, we asked trained undergraduate research assistants to make a subjective judgment about which toy infants preferred during the entire interaction (i.e., over the course of both the "looking-alone" and "interacting" phases) with each object pair. These results are available in Figure 2.10. A chi-squared test revealed some variability in the subjective judgments of objects that infants preferred during the object interaction task, $\chi^2(5) = 11.37$, $p = .04$. Again, while no one object stands out as having been overwhelmingly preferred, the Twisty Glasses do appear to have been less frequently preferred relative to the other objects. As with the other analyses, this figure separated by the identity of the objects that were paired is available in supplementary materials.

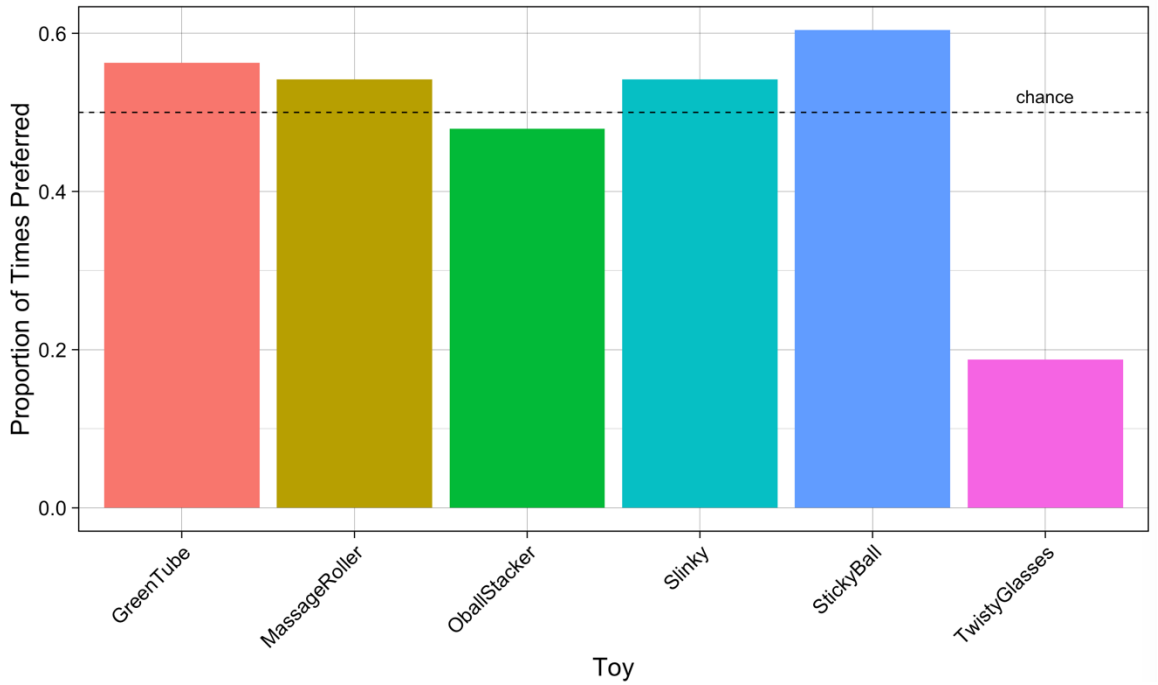


Figure 2.10. Proportion of times a given object was subjectively coded as being preferred by infants throughout the *looking* and *interacting* phases of the "infant interaction" task. The dashed line represents chance, which is .5 for any given object.

How familiar did caregivers believe these objects were to their infants?

Our final assessment of infants' familiarity with and preference for the set of objects featured in the pupillometry stimuli involved collecting parental report. Parents completed a survey asking if their infant had seen each object before coming into the lab and, if they had seen the object before, how likely it was that their infant knew what to do with that object. A chi-square test revealed that caregivers' ratings (e.g., yes, no, or maybe) did depend on the identity of a given toy, $\chi^2(10) = 55.17$, $p < .001$. These results can be viewed in Figure 2.11. For the majority of objects, parents indicated that their infant had never seen the object before coming to the lab. However, the Oball™ Stacker toy was nearly equally often rated as having been seen before coming in to the lab.

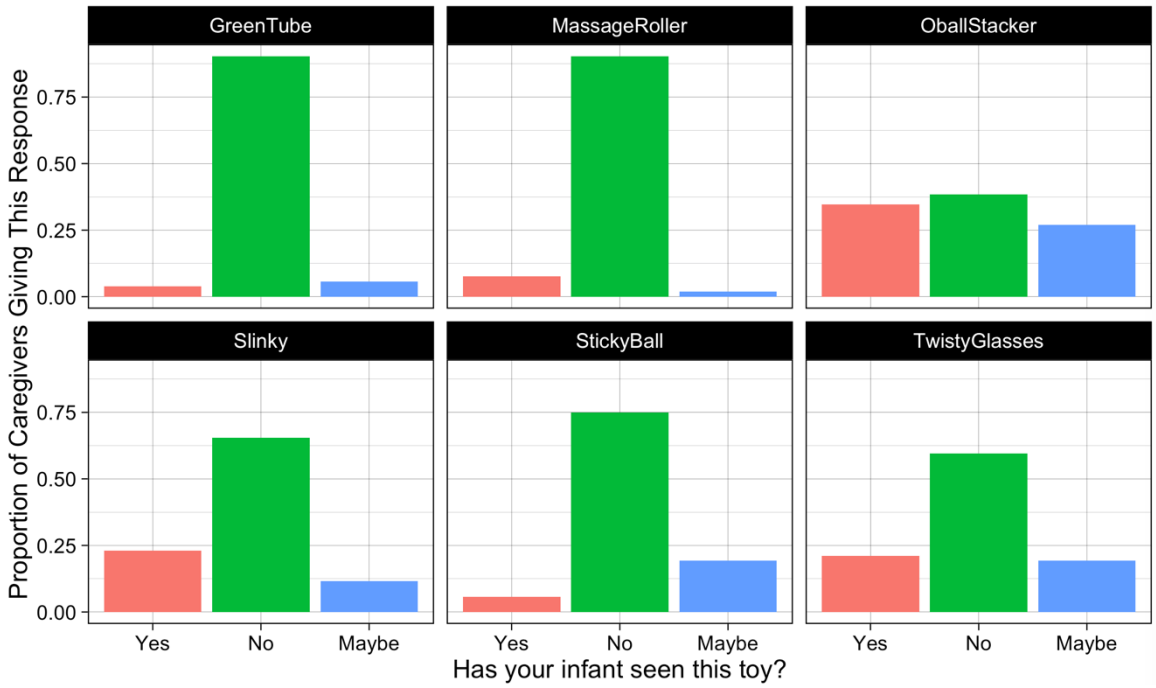


Figure 2.11. Proportion of caregivers who rated each object as "yes," "no," or "maybe" in response to the question of whether or not the infant had seen the object before coming in to the lab.

When caregivers reported that "yes" their infant had seen the object before coming in to the lab or that "maybe" their infant had seen the object, they were given the second

question asking how likely they believed it was, on a Likert scale of 1 to 5, that their infant came into the lab knowing what to do with that particular object (if caregivers reported that their infant had never seen the object before coming into the lab, they received a 0 on this measure). We asked this question in particular as there is some evidence that caregivers may be less likely to engage in motionese when their infant interaction partner already knows what to do with an object (Fukuyama et al., 2005). To explore this question, we ran a linear mixed effects model predicting caregivers' ratings from a fixed effect of object identity and a random intercept for subjects. Caregivers' ratings of how likely it was that infants came into the lab knowing what to do with the objects varied across the six objects, $F(5, 255) = 12.35$, $p < .001$. However, as can be seen in Figure 2.12, ratings were generally low. Even the object with the highest rating only had an average rating of 1.8 out of 5 on the Likert scale, suggesting that infants did not come in to the session with much knowledge of what to do with any of the six objects.

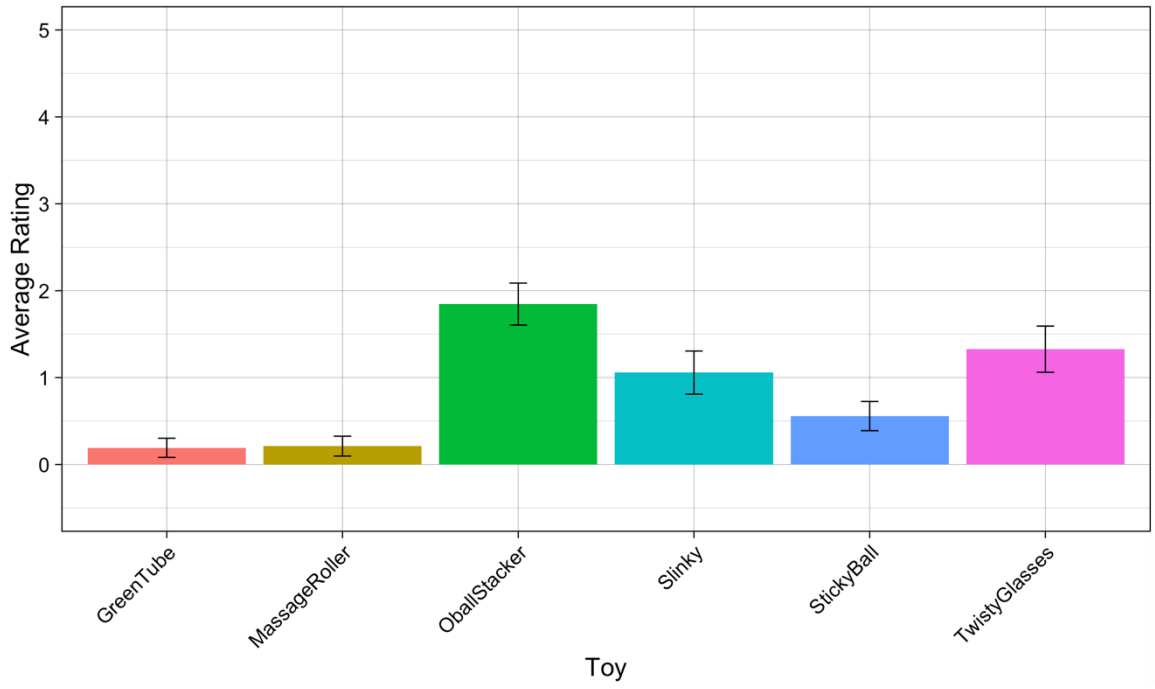


Figure 2.12. Average caregiver ratings (from 0 to 5) in response to the question "How likely is it that your infant came into the lab today knowing what to do with this object?" Objects are only included in this analysis if caregivers said that "yes" or "maybe" infants had seen the object before coming in to the lab. Error bars indicate +/- 1 SE.

To what extent did the videos depict motionese?

Our final set of analyses focused on trained research assistants' ratings of the twelve stimulus videos. Ratings on each of the eight dimensions of motionese (repetitiveness, rate, punctuation, range of motion, distance from partner, simplicity, interactiveness, and enthusiasm) ranged from 0 to 4. Higher values suggested more motionese[2] – for example, enthusiasm ratings ranged from 0 (very low enthusiasm) to 4 (very high enthusiasm). To explore differences in average overall ratings between infant- and adult-directed videos, we performed a mixed-effects regression predicting rating from a fixed effect of partner, a random slope and intercept for coder, and a random intercept for video. As predicted, average ratings for infant-directed demonstrations (M = 2.52, SD = 1.19) were significantly higher than ratings for adult-directed demonstrations (M = 1.77, SD = 1.05), β = 0.75, $t(9.43) = 4.95$, $p < .001$. Thus, on average, these ratings suggest that infant-directed demonstrations did indeed feature more characteristics of motionese.

To ensure that the age of the infant interaction partner did not influence the actor's demonstrations, we asked whether the above effect held when controlling for the age of the infant viewing the demonstration. In a model including the above fixed and random effects as well as an additional fixed effect of age and interaction between age

[2] One exception to this is the "rate" dimension which was coded 0: very slowly to 4: very fast. Here, we would expect demonstrations characteristic of motionese to be rated lower than adult-directed demonstrations. Ratings on this dimension were reverse coded to be in line with the other dimensions (i.e., higher ratings corresponded to more use of motionese).

and partner, interaction partner was still a significant predictor of rating, $\beta = 1.79$, $t(8) = 3.39$, $p = .01$, such that infant-directed demonstrations were rated higher in motionese on average. Age alone did not predict rating, $\beta = 0.001$, $t(8) = 0.134$, $p = .90$, nor was there a significant interaction between age and partner, $\beta = -0.003$, $t(8) = -2.04$, $p = .08$.

We next assessed the extent to which these ratings differed across the eight dimensions of motionese. We ran a linear mixed effects model nearly identical to the one describe above. However, in this model we additionally included fixed effects of dimension and the interaction between partner and dimension; we also removed the random slope for coder as including it caused issues with model convergence. Again, ratings differed across interaction partner, $\beta = 0.75$, $t(10) = 5.14$, $p < .001$, with infant-directed demonstrations rated higher in motionese. The omnibus test for dimension was also significant, $F(7, 450) = 10.90$, $p < .001$ as was the interaction between dimension and partner, $F(7, 450) = 17.31$, $p < .001$. As can be seen in Figure 2.13, ratings across the eight dimensions differed when the interaction partner was an adult versus an infant. The difference between infant- and adult-directed demonstrations was most pronounced for the distance from partner, enthusiasm, and interactiveness ratings ($ps < .001$). Ratings were slightly higher in infant- over adult-directed demonstrations on the simplicity, punctuation, and rate dimensions and slightly lower on the repetitiveness and range of motion dimensions, though these differences were small and did not reach statistical significance. These small differences between infant- and adult-directed demonstrations on some dimensions are likely due to restrictions we imposed to equate the infant- and adult-directed demonstrations. For example, selecting one action with an obvious boundary nearly eliminated the possibility that the infant- and adult-directed

demonstrations would differ on the dimension of repetitiveness. It is important to note that while motionese ratings were frequently lower for the adult-directed action, these ratings were not at floor. Thus, the infant- and adult-directed videos do not represent maximal extremes in motionese and instead reflect a more naturalistic contrast.

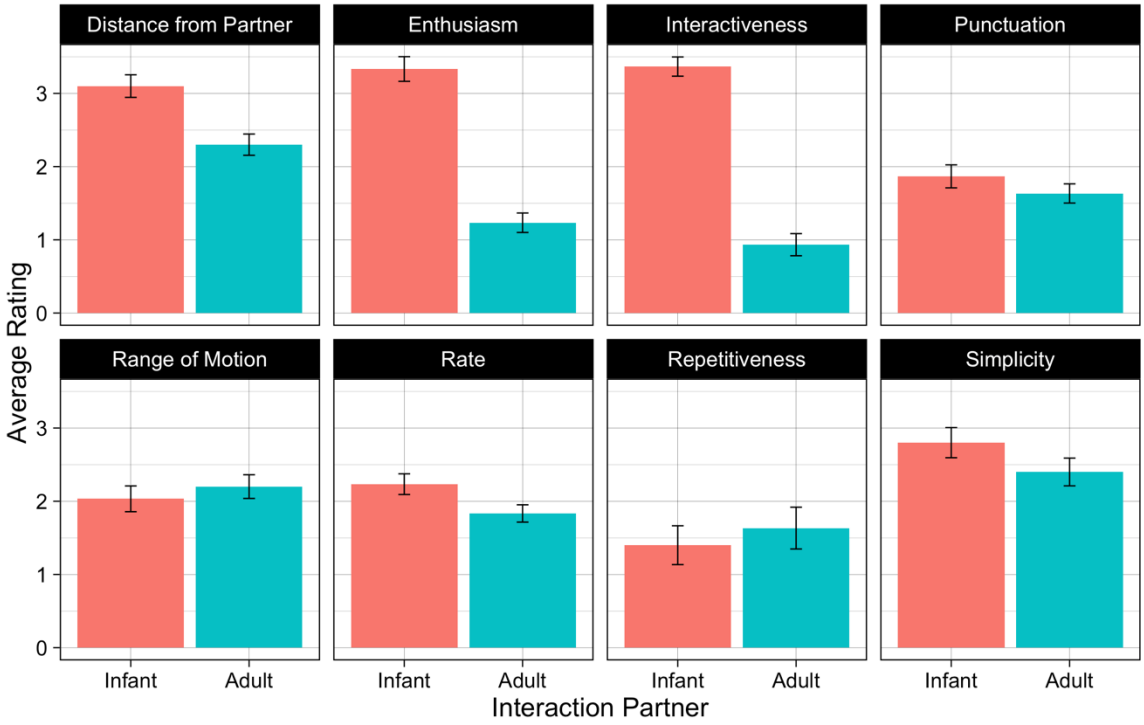


Figure 2.13. Average ratings for infant- and adult-directed demonstrations across each of the eight dimensions of motionese. Error bars indicate +/- 1 SE.

Discussion

To return to our goals for these analyses, they were to: (1) verify that the videos we chose for the pupillometry study did not systematically differ in ways that might influence our expected results (e.g., in luminance values), (2) collect baseline information about the salience of objects used in the videos, and (3) validate that actors' use of motionese varied along the expected dimensions in infant- versus adult-directed

sequences. While our luminance analyses alleviated some concerns about the potential for changes in luminance to influence our findings, we did find that luminance was lower at boundary regions of infant-directed action. These findings underscore the importance of controlling for luminance in our pupillometry analyses, especially those examining differences between pre-boundary, boundary, and post-boundary regions of activity. Our validation analyses of the location of action boundaries supported using expert judgments of action boundary locations, though we did find that the precise location of action boundaries were harder to identify for some videos. Regarding object salience and preference, there were no objects that were clearly overwhelmingly preferred or ignored across the analyses of infants looking to and interaction with objects nor through caregivers' ratings of infants' familiarity with objects. Finally, trained research assistants' coding of the twelve stimulus videos verified that infant-directed videos did exhibit more features of motionese than did adult-directed videos, at least on some dimensions.

Regarding luminance, the finding that luminance was lower at action boundaries for infant-directed demonstration went against our expectations. In our videos, luminance often varied with changes in the actor's position. If it is the case that infant-directed action serves to highlight action boundaries, we might expect a greater degree of motion change, and thus larger changes in luminance at boundary junctures within unfolding activity. One possible explanation is that infant-directed action highlights the region surrounding the action boundary, but does not place as much emphasis on the boundary itself, which would account for greater luminance values in pre- and post-boundary regions and smaller pixel values right at the boundary region. For example, a large, emphatic arm movement after a boundary might emphasize that a boundary has occurred,

but result in greater luminance at a post-boundary region. Additionally, while changes in luminance were lower at boundaries than pre- or post-boundary regions for the infant-directed action, the magnitude of the change was quite small. Therefore, the influence of purely luminance itself is unlikely to be consequential for our pupillometry results, though we will still control for luminance in all pupillometry analyses.

We found marked agreement on the location of action boundaries between experts and naïve participants, both when participants received information about the activity content that occurred at action boundaries and in the absence of this information. As expected, however, the correlation between participant and expert judgments was higher when participants received information about the activity occurring at the action boundary. Because our instructions were so specific in the first study (i.e., we told participants exactly what boundary to look for), we might have expected even higher correlations between participant and expert judgments. Research assistants running naïve adult participants through the task noted that some individuals were much less precise when instructed to use the arrow keys to find the location of the action boundary. In line with this observation, we confirmed that participants' judgments of the boundary were more likely to be late than early, suggesting that participants may be pausing the video after seeing the boundary but failing to use the arrow keys to return back to the precise boundary moment. This lack of precision across naïve participants was present in both versions of the boundary agreement study (with and without precise instructions about the content of the action boundary) and likely explains much of the reason for lower than expected point-biserial correlations.

Additionally, it appeared to be somewhat more challenging to identify the action boundary location in some of the videos. In particular, there was more variability in participants' judgments of the boundary location for the infant-directed demonstration of the Green Tube. In this activity sequence, participants were asked to find the moment at which the actor finishes pulling the Green Tube. The lack of agreement on the location of the action boundary likely occurred because the actor stretches out the tube very slowly, and it is challenging to know exactly when she has stopped pulling.

In the tasks assessing infants' baseline preference and familiarity with objects used in the pupillometry videos, infants' looking, infants' interaction with objects, and caregivers' object ratings did not reveal a single object or set of objects that were overwhelmingly preferred by or familiar to infants. However, there were some consistencies across tasks suggesting that some objects might be slightly more or less preferred than others. One of the findings that emerged is a lack of preference for the Twisty Glasses. This was the object infants interacted with least in the twenty-second "interacting" task and occurred least often as infants' preferred object (as subjectively coded by RAs), though these measures are likely to be highly correlated - time spent interacting with an object is likely to have influenced RAs' judgments about infants' preferred object. Another relevant finding is that infants' first look was most often to the Oball™ Stacker toy, and this toy also received the highest rating in response to the question of whether infants had come into the session knowing what to do with the object. Perhaps infants looked to this toy first because it was already familiar to them. However, the average rating for this object was only a 1.8 out of 5, so despite it being the "most" familiar object by this measure, the rating was very low. Still, taking these

findings into consideration, it will be important to account for object-related variance in the pupillometry analyses (i.e., by including a random effect of video).

In examining trained research assistants' motionese ratings we found that infant-directed demonstrations were higher in motionese on average. This is in line with our expectations, as these videos were chosen because they subjectively appeared to be high in motionese. While this effect held when in a model that included motionese dimension, dimension was also a significant predictor of motionese rating. In particular, ratings for distance from partner, enthusiasm, and interactiveness were significantly higher (i.e., more motionese) in the infant-directed demonstrations. The fact that infant- and adult-directed demonstrations differed less in motionese on some dimensions could be due to constraints we placed on the videos in our attempt to equate infant- and adult-directed action sequences. For example, as described earlier, using a single action with one boundary eliminates much of the opportunity for repetition. Additionally, equating the demonstrations for length in the context of a matched activity is also likely to decrease any possibility for differences in rate across the paired videos.

Overall, the resulting set of twelve infant- and adult-directed action videos seem appropriate for exploring the effects of motionese on infants' processing of unfolding activity using the pupillometry methodology. In addition to providing stimuli for the pupillometry experiment, this corpus provides a rich set of video data with which to explore the nuances of adults' modifications to infant-directed action in future research.

CHAPTER III

USING PUPILLOMETRY TO ASSESS THE INFLUENCE OF MOTIONESE ON INFANTS' PROCESSING OF DYNAMIC ACTIVITY

Introduction

The overarching goal of this dissertation was to explore the extent to which caregivers' modifications to infant-directed action (e.g., "motionese") influence infants' processing of activity as it unfolds across time. As has been described previously (both in prior work in the domain of motionese and in the work described here in Chapter II), infant- and adult-directed action differ across a number of dimensions including repetition, range of motion, interactiveness, exaggeratedness, and a variety of other characteristics. Motionese demonstrations engage infants' overall attention (e.g., Brand & Shallcross, 2008) and promote imitation (e.g., Williamson & Brand, 2014). However, less is known about specific ways in which modifications to infant-directed action impact infants' online processing of unfolding activity. Information of this kind provides insight into infants' action processing propensities and how motionese may dovetail with such propensities to support infants' action processing. In particular, previous research indicates that the ability to rapidly and efficiently extract segmental structure from activity as it unfolds across time is important for understanding, remembering, and performing action (e.g., Kurby & Zacks, 2008; Levine et al., 2018). We hypothesized that a key benefit of motionese is that it helps infants identify and attend to this internal segmental structure of dynamic human action.

To return to the three objectives outlined in Chapter I, we explored the extent to which (1) motionese enhances infants' overall attention to action, (2) infants selectively attend to action boundaries in continuous activity sequences, and (3) motionese influences infants' attention to the structure of unfolding activity. To investigate these questions, infants viewed a series of videos depicting adults demonstrating brief action sequences to their own infant and to an adult interaction partner, described in detail in Chapter II. Each of these videos contained one major action boundary. Using the SIPR system (Bala et al., 2016), as infants viewed these videos, their pupil diameter was recorded via a Raspberry Pi NoIR camera connected to a Raspberry Pi computer (this system is described in further detail below). Because this methodology enabled us to monitor infants' attention to, and engagement with, streaming visual information, as indexed by pupil dilation, we were able to compare infants' processing of infant- and adult-directed action sequences as they dynamically unfolded across time. We anticipated systematic differences in infants' processing of motionese versus adult-directed action.

First, we expected to replicate previous research suggesting that infants preferentially attend to motionese over adult-directed action in two different ways. Using pupillometry, evidence of a replication of this result would appear as enhanced processing (indexed by greater tonic pupil dilation) for infant- over adult-directed demonstrations. We additionally attempted a direct replication (i.e., using looking time as a dependent measure rather than pupil dilation) of previous studies demonstrating that infants prefer to look at motionese over adult-directed action. To carry out this replication, we simply analyzed the amount of time (as indexed by the Raspberry Pi recording) that infants spent looking to infant- versus adult-directed demonstrations. We

predicted that the time infants spent looking to motionese would be significantly longer than to adult-directed action, as has been observed in previous research in the domain of motionese (e.g., Brand & Shallcross, 2008).

Regarding the extent to which infants selectively attend to action boundaries, some evidence exists that infants are sensitive to the structure of unfolding activity, and preferentially attend to boundaries (e.g., Levine et al., 2018; Baldwin et al., 2001; Stahl et al., 2014; Hespos et al., 2009, 2010). Therefore, we expected to find enhanced processing (indexed by increases in phasic pupil dilation) at action boundaries over non-boundary regions across both motionese and adult-directed activity sequences.

Finally, if motionese indeed assists infants in finding structure within unfolding activity, the magnitude of increases in pupil dilation at action boundaries should be larger when action is presented using motionese relative to when action lacks such modifications. In sum, we explored both tonic and phasic components of infants' pupil dilation. We predicted that overall (tonic) pupil dilation would be larger for infant-directed action and that, on top of this tonic effect, the phasic effect of increased pupil dilation to action boundaries (relative to pre- or post-boundary regions) would be magnified for motionese relative to adult-directed action. Schluroff (1983) demonstrated a similar effect in language processing research, finding phasic responses (brief increases in pupil size) to word onsets on top of variation in overall average (tonic) pupil size in response to sentence difficulty.

In addition to exploring infants' pupillary response to motionese versus adult-directed action, we explored the extent to which infants chose to interact with an object that had been demonstrated in either a motionese versus an adult-directed format. We

predicted that infants would be more interested in, and more likely to choose to interact with, objects that had been demonstrated using features of motionese.

Method

Participants

Twenty-eight infants ranging from 9 to 12 months (14 females; Mean = 314 days; SD = 34.7 days) and their caregivers participated in the pupillometry study. One infant was immediately excluded due to serious medical issues at birth. Our reasons for selecting this age range were similar to that of the corpus creation project outlined in Chapter II. Specifically, infants at this age are likely to receive and be sensitive to motionese input (e.g., Brand et al., 2002; Brand & Shallcross, 2008). By about 9 months of age, infants additionally begin to attend to relationships between people and objects rather than simply attending to people or objects alone (i.e., "secondary intersubjectivity;" Carpenter et al., 1999; Baldwin & Kosie, 2019; Trevarthen, 1977; Trevarthen & Hubley, 1978; Rochat et al., 2009; Bakeman & Adamson, 1984). Infants in this age range have additionally acquired the motor skills necessary to explore toys themselves (e.g., Lockman & McHale, 1989; Baldwin et al., 1993; Kimmerle et al., 1995).

However, while we extended our age range to 18 months for creation of the video corpus, we chose to test only infants from 9-12-months of age in the pupillometry study. We chose to restrict the age range for two main reasons. First, the more restricted age range allowed us to better equate motor skills across infants. Recall that our goals in the

corpus creation were to characterize adults' modifications to infant-directed action and generate a set of video stimuli for the pupillometry study that contained both infant- and adult-directed action sequences, while maximizing the difference between infant- and adult-directed demonstrations. In contrast, in the pupillometry study, we were specifically interested in infants' processing of action. Given evidence that infants' own action experience influences their processing of others' action (e.g., Sommerville, Woodward, & Needham, 2005), it seemed important to control infants' action experience to at least some degree across participants in the current study; constraining the age range was one way to accomplish such control. Additionally, we expected infants' knowledge about objects and actions to impact their pupil dilation response. Thus, a more homogeneous sample from a restricted age range was likely to decrease noise and variability in observed patterns of pupil dilation.

Families from the local Eugene, OR community were recruited to participate through the University of Oregon Psychology Department's developmental database. Race/ethnicity of caregivers and infants was representative of the general Eugene, OR community. All participants (28; 100%) identified as White, 4 participants (14%) additionally identified as Hispanic, 1 participant (4%) additionally identified as Asian, and 1 other participant (4%) additionally identified as Indian or South Asian (caregivers were asked to select all races that applied). To assess socioeconomic status, each family provided information about maternal education (as mentioned in Chapter II, maternal education is a proxy for SES that tends to be predictive of developmental outcomes; e.g., Gottfried et al., 2003; Noble et al., 2007; Liaw & Brooks-Gunn, 1994). As in Chapter II, mothers in our sample generally reported high educational achievement, with 38%

reporting some level of graduate training (see Table 3.1 for detailed information). After participating, all families received their choice of either a t-shirt or a children's book as a thank you gift.

Table 3.1.

Highest level of maternal education across caregivers in the pupillometry study. We report both the number of caregivers having achieved each level of maternal education as well as the proportion of the sample that this number represents. One caregiver did not provide the gender of caregivers, thus maternal education could not be determined, and they are not included in this summary.

Maternal Education	**Number of caregivers**	**Proportion of sample**
High School	1	3.6%
Some College	2	7.1%
Associate's Degree	0	0%
Bachelor's Degree	13	46%
Master's Degree	5	17.9%
Doctoral Degree	6	21.4%

Apparatus

Infants were seated in a car seat approximately 82cm from a black floor-to-ceiling curtain, in front of which was a 58cm wide-screen monitor that presented stimuli at a size of 1920 x 1080 pixels. Infants were strapped into the car seat by the caregiver, and straps were pulled snug to secure infants into the seat. Additionally, the car seat contained padding on either side of the infant's head, decreasing the amount of head movement that was possible. Infants' movement was not otherwise restricted. Pupil dilation was digitally recorded via a Raspbery Pi NoIR camera (infrared camera) placed approximately 38cm from the infant's eyes, just out of reach. Video from the camera was recorded to a Raspberry Pi single-board computer at a rate of 30 frames per second. Two small infrared lights were placed on either side of the Pi camera and a third, larger, infrared light was placed immediately to the left of the Pi camera. These lights helped to illuminate the

infant's face and make the pupils more readily detectible on the resulting video recording. A second SONY video camera was placed above the monitor and zoomed in to gain a close view of the infant's face. The video file to which this camera recorded was synchronized with the video being played to the infant, resulting in a recording of the infant's face that also depicted, in the top left corner, what the infant was seeing. This second video was used for hand coding infants' looking throughout the pupillometry session. Figure 3.1 depicts the experimental setup.

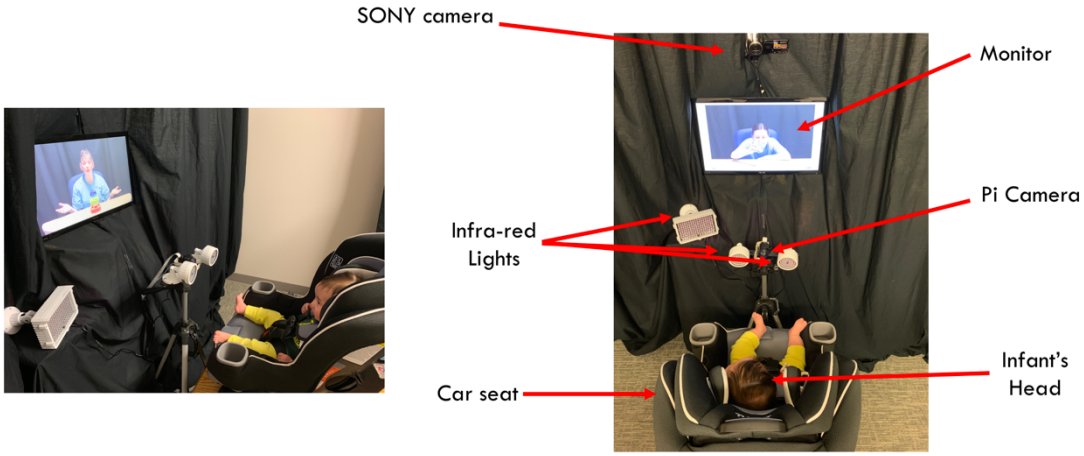


Figure 3.1. Experimental setup. Infants were seated in a car seat facing a computer monitor. A Raspberry Pi NoIR camera, placed approximately 38cm from infants, recorded their pupillary response to a Raspberry Pi single-board computer. A SONY camera recorded infants' looking to the screen.

Design

As described in detail in the previous chapter, we generated a set of twelve possible videos that could be shown to infants. These twelve videos consisted of six video pairs, each pair depicting the same mother performing the same actions on the same object for their own infant or for an adult research assistant. All videos contained

one major action boundary that occurred at approximately the same location across the infant- and adult-directed versions of the same activity. To avoid effects of familiarity with actors and objects, each infant saw only one video from each pair; that is, they saw *either* the infant- and adult-directed version of each activity. As a result, each infant viewed six unique videos, three infant-directed and three adult-directed. A set of six videos (three infant-directed, three adult-directed) constituted one "block." Infants viewed up to six total blocks.

Because we could not fully counter-balance the videos given our expected sample size, we opted to randomly choose two groups of videos (three infant-directed and three adult-directed in each) and assign an equal number of infants to each group. Based on random assignment, one group of infants viewed the infant-directed versions of the Slinky, Sticky Ball, and Massage Roller and the adult-directed version of the Green Tube, Oball™ Stacker, and Twisty Glasses. The second group saw the opposite (adult-directed Slinky, Sticky Ball, and Massage Roller and infant-directed Green Tube, Oball™ Stacker, and Twisty Glasses). To ensure that infants never saw more than two infant- or adult-directed videos in a row, the video presentation for half of the infants in each group followed one pattern (Blocks 1, 3, and 5: infant-directed, adult-directed, adult-directed, infant-directed, infant-directed, adult-directed; Blocks 2, 4, and 6: adult-directed, infant-directed, infant-directed, adult-directed, adult-directed, infant-directed) and the presentation for the other half of the infants followed the opposite pattern (Block 1, 3, and 5: adult-directed, infant-directed, infant-directed, adult-directed, adult-directed, infant-directed; Blocks 2, 4, and 6: infant-directed, adult-directed, adult-directed, infant-directed, infant-directed, adult-directed). The actual videos that occurred in each of these

positions were randomly assigned from the infant- and adult-directed versions in that infant's group[3].

Stimulus presentation was programmed in PsychoPy (Peirce, 2007). As depicted in Figure 3.2, all blocks started with a brief video of a laughing baby as an attention-getter to help infants orient to the monitor, which played for eight seconds. After the laughing baby attention-getter, a set of moving concentric circles played for three seconds as infants heard a chime sound. The laughing baby and chimes stimuli were acquired via publicly shared materials from the ManyBabies study of infant-directed speech preference (ManyBabies Consortium, under revision). While the laughing baby attention-getter was only played at the start of each block, the circle and chimes attention-getter was played before each video. For three seconds at the beginning of each video, infants were presented with a grey screen signaling the start of the trial. A secondary goal of the grey screen was to match the luminance of the first frame of the video. Due to an inadvertent change in luminance of the grey screen during stimulus creation, this goal was not satisfied. However, because we were not specifically interested in infants' PDR to content immediately following the grey screen, this issue was not problematic for interpretation of our results (and is thus explained in further detail in supplementary materials). After the grey screen, infants were then presented with a three-second still image depicting the first frame of the action sequence. The still image was included to allow infants' pupils to adapt to both the luminance and the characteristics (e.g., featured actor and object) of the visual scene that would be viewed in the upcoming video. After

[3] This randomization was not fully successful. It turned out that some pairs were more common than others (e.g., the Green Tube was presented before or after the Slinky much more frequently than it was presented before or after the Massage Roller). However, our incorporation of a random effect for videos in mixed effects models provides confidence that this issue did not undercut interpretation of the findings.

the three-second still, the action sequence began to play silently at a standard rate of 30 frames per second[4]. Upon completion of a trial, the infant again heard the chimes while viewing the concentric circles, and then the next trial started with a grey screen followed by a still frame. Once infants had completed their six unique trials the laughing baby played again, starting the next block. This repeated for a total of six blocks or until the infant became too fussy to continue.

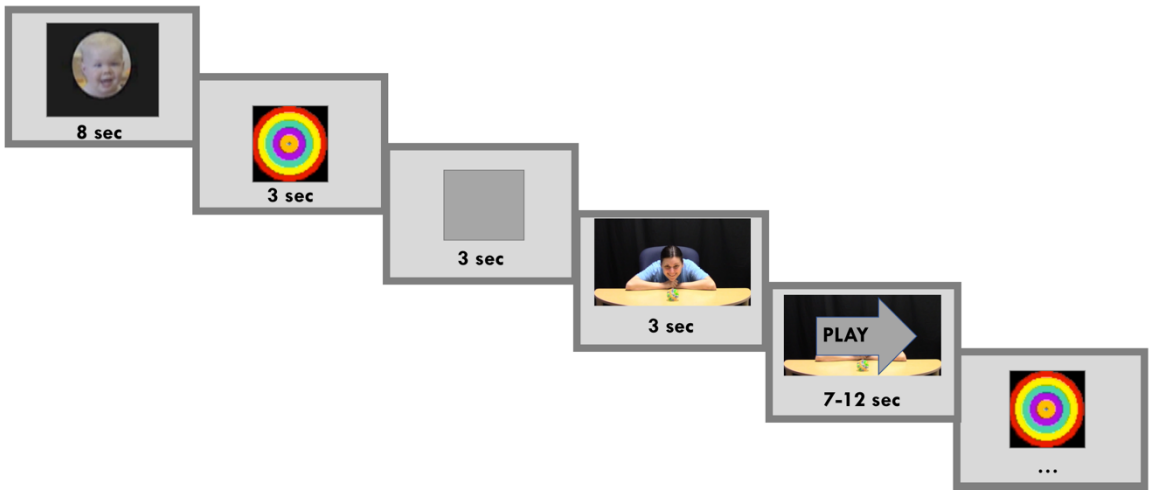


Figure 3.2. This figure depicts the structure of the pupillometry experiment. Each block started with a laughing baby attention-getter. Infants then heard chimes while viewing concentric circles. At the start of each trial, infants viewed a grey screen for 3s and then the first frame of the video for 3 seconds. After this, the video silently played in its entirety. Infants viewed the circles/chimes between each trial. Six trials constituted a single block, and each block repeated a maximum of six times.

[4] In the process of data analysis we learned that stimulus files frequently took variable amounts of time to load (in the range of approximately one second). While files were loading infants simply viewed a black screen. Because this loading delay occurred prior to the grey screen, it is very unlikely to be problematic for interpretation of our results. However, details regarding how this video lag was dealt with in the process of data analysis are included in supplementary materials.

Procedure

Caregivers were seated shoulder-to-shoulder with the infant, but facing away from the monitor. This setup allowed infants to see the caregiver should they look over, but avoided the possibility that infants would be influenced by any caregiver reaction to the stimuli. Caregivers were asked to remain facing away from the monitor and not to interact with the infant. We requested that, if the infant started to fuss, they simply put their hand on the infant. However, if at any point they wanted to take a break or stop the experiment, they should feel free to let us know and we would stop immediately. Once the caregiver and infant were seated, the experimenter adjusted the focus of the Raspberry Pi NoIR camera to ensure a clear picture of the infant's pupil. She then went behind the curtain and began the pupillometry session. If infants completed the entire set of six blocks, this part of the session lasted approximately 12 minutes.

After the pupillometry portion of the session, caregivers transferred the infant to a highchair in the center of the room and a video camera was placed in front of the highchair, recording the infant's actions. The experimenter brought out the six objects infants had viewed in the pupillometry portion of the study. She presented these objects to the infant in three pairs, each pair containing one object the infant had viewed in a motionese demonstration and another that the infant had viewed in an adult-directed demonstration. The order of object presentation matched the order in which infants viewed the objects – for example, the first pair consisted of the first two objects the infant had viewed in Block 1. Because we specified the order in which infants viewed motionese and adult-directed demonstrations, the objects could be paired based on trials 1 and 2, 3 and 4, and 5 and 6. For two members of each pair, one had always been featured

in an adult-directed demonstration and the other always in an infant-directed demonstration. We additionally randomized whether infants were presented with the object they had viewed in the infant-directed demonstration on the right or left side during the interaction task. At the beginning of this task, the experimenter started a timer with a ticking clock sound and placed the first pair of objects on a serving tray out of the infant's sight. As in the corpus creation task, the experimenter held up the tray so that the infant could see, but not reach, the objects and said, *"Look what I have!"* She held the tray still for three seconds before saying *"Here you go!"* and placing the serving tray on top of the infant's highchair tray. Infants were then allowed to interact with the objects for twenty more seconds.

Next, caregivers and infants engaged in free play with a different set of six new objects (i.e., toys that had not been featured in the video stimuli): three familiar and three novel. At the end of the free play period, caregivers completed a survey about the extent to which these objects were likely to have been familiar to their infants prior to coming in for the study. These data will be analyzed separately and are not part of this dissertation work, thus they are not discussed further.

Caregivers also completed a basic demographics questionnaire and the Infant Behavior Questionnaire (IBQ; Rothbart, 1981; Putnam, Helbig, Gartstein, Rothbart, & Leerkes, 2014). The IBQ measure of infant temperament asked caregivers to report on the frequency with which the infant had engaged in certain behaviors over the course of the prior week. After completing these questionnaires, caregivers were presented with a Databrary (Databrary, 2012) consent form and asked if they would be willing to allow us to share their videos with other researchers. Before coming in to the lab, caregivers had

been emailed and asked to complete the MacArthur Communicative Development Inventory (MCDI) administered through Web CDI (Fenson et al., 1994, 2007). The IBQ and MCDI data will be analyzed for future research, results are not included in this dissertation.

Inclusion Criteria

As mentioned previously, each block of videos contained six unique trials, each trial depicting a unique motionese or adult-directed action sequence. Infants had the opportunity to view six total blocks of videos, each containing the same six trials (i.e., activity sequences). The experimenter, blind to which video the infant was viewing, noted times at which the infant was fussy. An infant was considered to be fussy if they were crying without pause. Infants occasionally fussed for a block or two and then became re-engaged with the video. In these cases infants were considered "fussy" only for those few blocks. Some infants did not re-engage after fussing, thus video presentation ended early and the maximum possible number of blocks they contributed was lower. Trials were considered unusable if the infant was fussy (as coded by the experimenter) and/or if the Matlab program indicated that the infant was not looking at the screen for at least 50% of the trial. The number of unusable trials was approximately equal for motionese and adult-directed trials. An entire block (i.e., one presentation of the six videos) was dropped from analysis if an infant's data were unusable on more than 50% of trials within the block. All infants in the current study contributed at least one block of data, thus none were completely excluded from analyses. In total, 696 trials across 27 infants were included in the pupillometry analyses. The median number of

trials contributed by each infant was 29 (out of 36 total possible trials). A table specifying the total number of full blocks, partial blocks, and trials contributed by each participant is available in supplementary materials.

Coding

Infant Gaze. Because the pupillometry methodology implemented in this research was new, we opted to hand-code infants' looking for a subset of videos as validation of how well the automatic pupil-detection algorithm detected moments at which the infant was looking to the screen. For 25% of the participants, videos collected from the SONY camera were hand-coded for infants' gaze during presentation of the pupillometry stimuli. A trained research assistant coded these videos in two passes. First, she covered the top right corner of the screen (in which the content viewed by the infant was visible) and coded whether or not the infant was looking at the computer monitor; thus, her coding was blind to the stimuli infants were viewing. Before beginning to code, she watched parts of the video in real time to get a sense of what the infant's pattern of looking was like. She then went through the video frame-by-frame and coded, in Datavyu (Datavyu Team, 2014), whether infants were looking at the monitor. This Datavyu coding resulted in the precise timing of the onset and offset of infants' looks to the monitor. In a second pass through, the coder uncovered the corner of the screen depicting the infants' visual input. She proceeded to code the onset and offset of each part of the pupillometry stimuli (e.g., laughing baby, chimes, grey screen, still frame, activity sequence). These two sets of coding could be combined to align infants' hand-coded looking patterns with automated analyses produced by running the Pi videos through a

Matlab pupil detection script (Matlab, 2019; described below). Before coding these videos, the coder underwent training in this procedure. To "pass" this training, she had to be reliable with at least 90% of an expert's frame-by-frame judgments on a set of training videos.

Infants' object-interaction task. All infants had the opportunity to interact with the six objects they had viewed in the videos, presented in three pairs. As in the corpus creation study, infants first viewed the objects for three seconds and then interacted with the objects for twenty seconds. A trained research assistant coded these videos in two passes (again in Datavyu; Datavyu Team, 2014). This coder did not have any knowledge of which videos the infant had viewed and thus was blind to the identity of which object in each pair had been featured in motionese versus adult-directed demonstrations. First, she coded the item that the infant looked to when initially presented with the object pair, pausing the video immediately after the researcher said *"Look what I have!"* and specifying the object to which the infant was looking. Next, she coded the duration of infants' looking to each of the objects during the three-second looking-alone phase and their interest in the object during the twenty-second interacting phase. As in the corpus creation study, coders were told that interest in an object could include looking, touching, or manipulating an object, but to keep in mind that infants may be touching an object they are not interested in or interacting with. During this pass, coders additionally made a subjective judgment about which object the infant "preferred" across the looking-alone and interacting phases. Again, instructions are available on the OSF page associated with this dissertation (http://osf.io/8mzhf). To assess reliability, 20% of videos were coded by a second research assistant for all of these passes. Coders agreed on the object of infants'

first look on 83% of trials, Cronbach's alpha was 0.93 for the duration of infants' looking and 0.98 for the duration of infants' interacting with objects, and coders agreed on the "preferred" object on 86% of trials.

Data acquisition

We recorded a separate video for each infant via the Raspberry Pi NoIR camera and Raspberry Pi computer. Each video was run through a Matlab (Matlab, 2019) pupil detection program designed to advance frame-by-frame through the video, find circles, and measure their diameter. First, the Matlab program read in all of the frames from the video file and stored them in memory. At this time, we synchronized the videos collected from the Pi Camera and the stimulus presentation in PsychoPy (Pierce, 2007). To do this, we had programmed PsychoPy to flash a UV light in the corner of the camera's visual field at the exact moment the grey frame appeared on the screen (i.e., at the start of each trial). Thus, to synchronize the videos, we used the Matlab program to detect flashes and calculate the difference between that moment and presentation of the grey screen in the data recorded by the PsychoPy program. We were then able to align the infants' pupil dilation with the stimulus that was being presented by PsychoPy at any given time throughout the experiment. As a timing check, we were also able to calculate an offset between the time at which the UV light flashed (identified by Matlab) and the start of each trial (reported in the PsychoPy output) for all trials in the video. These values were consistent across trials for all participants, suggesting that the timing of Pi Camera recording and PsychoPy stimulus presentation software were well aligned.

We next defined a number of additional parameters, specific to each video, that enabled the Matlab (Matlab, 2019) program to detect and measure circles. The first step in calculating pupil diameter was to manually examine the video to determine which of the infants' eyes was visible most frequently throughout the video. We defined that eye as the one that the program should detect and for which it should measure pupil diameter. If, for example, we chose the left eye, the program would calculate the diameter of the circle that was closest to the left side of the image. While this usually meant that the diameter corresponded to the left eye, occasionally infants moved their heads and the left eye was not visible on the screen. If the right eye was visible at these moments, it would become the left-most circle on the screen and thus the diameter of the right pupil would be calculated (note that right and left eye pupil dimeter are strongly correlated; Jackson & Sirois, 2009; Sirois & Jackson, 2012). We next set a number of metrics that enabled the Matlab program to detect and measure pupils. To set these metrics, we selected one frame at random for which the infant's pupil was clearly visible. We next pushed this frame to threshold so that everything but the pupils faded to white. Specifically, for each participant, we set a threshold for how dark a pixel had to be to remain black, and all pixels in the frame that were not at least as dark as that value became white. After pushing the image to threshold, we used the Matlab Data Cursor to measure infants' pupil size in the randomly-selected sample image, and we referenced this measurement to set limits for the size of circles to be detected (setting limits too low would allow for things like the infants' nostrils to be detected as circles, while setting limits too high would, in some cases, include things like the infant's hair being considered a circle). We additionally set a sensitivity metric; this metric specified how precise the shape and size

of a potential pupil image on a given frame had to be to consider that shape a circle, and thus to calculate a pupil diameter. With these metrics, we detected pupils for one image and plotted, over the image, a red circle that indicated the circles that had been detected and measured. This allowed us to visually assess how well the metrics that were set corresponded to the pupils visible in the image. After setting these metrics for one frame in the video sequence, we randomly selected ten additional frames to validate that these metrics were able to detect pupils throughout the video sequence. Again, for each of these frames, we plotted the circles that were detected over an image of the infants' pupils and visually examined. The values chosen for these settings for each participant are available on the OSF page associated with this dissertation (http://osf.io/8mzhf). Once these metrics were set and verified, the program automatically used them to calculate pupil diameter for each frame in the video recorded by the Raspberry Pi camera.

We next turned to decisions regarding data interpolation, z-scoring, filtering, and baseline correcting. This presented a challenge as, in the field of infant pupillometry research, there is a lack of consistency across studies in the implementation of preprocessing steps (Geller, Winn, Mahr, & Mirman, 2019; Hepach & Westermann, 2016; Mathôt, Fabius, Van Heusden, & Stigchel, 2018). Thus, rather than follow one specified procedure for pupillometry analyses, we examined prior research to make informed decisions about the preprocessing steps that were most appropriate for the current study.

Following previous research using this pupillometry system (Bala et al., 2016) as well as that of other experts in pupillometry research (e.g., Unsworth & Robinson, 2015; Miller, Gross, & Unsworth, 2019). As in this previous research, we opted not to

interpolate missing values to preserve the original data to the extent possible in the current analyses. We did, however, engage in a number of data manipulation procedures in an effort to render the data more interpretable and comparable across subjects. First, we z-scored pupil size measurements for each participant. To calculate z-scores, we included all relevant frames for each participant (i.e., data from the grey screen, still frame, and video across all blocks and trials but ignoring responses to the attention-getting stimuli) and used these same z-scored data across all analyses. Specifically, we calculated the mean and standard deviation of pupil size for each participant (across all blocks and trials), subtracted the individual's mean from their pupil diameter at each frame, and divided this value by that individual's standard deviation. Z-scoring was done for the following reasons: (1) the Matlab program records pupil diameter in pixel size, which is dependent on features of the Pi video (e.g., the degree of zoom on the infants pupil) thus z-scoring made the pupil size measurements more comparable across participants, (2) z-scoring both pupil diameter and luminance makes these measurements more interpretable and more easily comparable as well, and (3) z-scoring controls for individual baseline pupil size differences across subjects, while (4) preserving within-participant pupil diameter differences across motionese and adult-directed action.

After z-scoring, the raw pupil values were filtered to eliminate random fluctuations in the data. While there are multiple possible filters that can be used to smooth pupillometry data (see Hepach & Westermann, 2016 for a review in infancy research), we opted to use a hanning filter with a standard window size of 11 frames. We chose this filter for several reasons. For one, the hanning filter uses a moving average, which is one of the common ways of filtering data in pupillometry research and is among

those suggested by the creators of R packages for analyzing pupillometry data (e.g., Hepach & Westermann, 2016; Geller et al., 2019). Additionally, the hanning filter can handle missing data, enabling us to perform pupillometry analysis without first interpolating missing values due to blinks or "look aways." Finally, a visual comparison of filtered and unfiltered data suggested that the hanning filter would appropriately preserve effects of interest while removing extreme values. The hanning filter uses a weighted moving average by generating a normal distribution of weights centered on the frame being filtered and encompassing the surrounding 10 frames (when the window is set at 11, which is the standard, recommended window in pupillometry research). Because of this distribution of weights, the frame of interest has the largest influence on the filtered pupil value, the frames on either side have the next largest influence, and the amount of influence decreases until the distribution covers 11 total frames. Frames outside this window do not contribute to the estimate of pupil size. The z-scored, filtered data are referred to simply as "pupil size" for the remainder of this manuscript.

After filtering the data, we generated baseline-corrected values for each participant on each trial. Our measurement of baseline was the average pupil diameter in the one second region before onset of the video, calculated separately on each trial. This baseline was chosen for two reasons. First, on viewing infants' pupil diameter to the grey screen, still frame, and start of the video, it appeared that there was a large luminance effect when the grey screen changed to the still frame. This luminance effect appeared to take about two seconds (of the three-second still frame) to begin to recover. Thus, the final one second of the still frame seemed to be the most appropriate baseline

measurement[5]. Secondly, while there is no real consensus across pupillometry research in how to choose a baseline value, the one second before stimulus onset baseline has been used in a number of infant pupillometry studies (e.g., Geangu et al., 2011; Hepach & Westermann, 2013; Morita et al., 2012, Nuske et al., 2015). In line with methods used by Tanaka and colleagues (in preparation), we opted to control for baseline in analyses via covariation rather than correct for baseline via subtracting or dividing pupil sizes by baseline pupil size. Both subtracting or dividing pupil size by a baseline value have the disadvantage that, when infants' pupils are large at baseline, the degree of possible change as they view the videos would be diminished and therefore it would be harder to detect stimulus effects.

Results

Our goals in the current analyses were to (1) examine the effect of motionese on infants' looking and pupil diameter, (2) explore the extent to which infants' pupil diameter is indicative of action segmentation, and (3) investigate the influence of motionese on infants' action segmentation. We additionally assessed infants' interest in the objects they had viewed. Before turning to these analyses, however, we report the results of a set of "validity checks" we performed to examine whether the data reflected expected patterns of looking and PDR. We also report the results of hand-coded validation of the automated pupillometry procedure.

[5] While we used a one second baseline for the analyses reported here, we additionally performed all analyses with a 3s baseline and without covarying baseline at all. The general patterns and significance of the results did not change for either of these analyses.

For analyses estimating linear mixed-effects models, we used the *lme4* package (Bates et al., 2015) in *R* (R Core Team, 2018) with type III sums of squares (set using the *afex* package; Singmann, Bolker, Westfall, & Aust, 2017). Significance for these models was assessed using the *lmerTest* package (Kuznetsova et al., 2015; Luke, 2017) with Satterthwaite's approximation for degrees of freedom. All analyses began with a maximal model, including random slopes and intercepts for subjects and videos (Barr et al., 2017). However, this fully random model rarely converged in current analyses. If models did not converge, we first removed random slopes for videos and then random slopes for subjects, keeping random intercepts. We have specified the exact fixed and random effects structure used for each model below. For analyses requiring pairwise comparisons, we used the *lsmeans* package with a Bonferroni correction for multiple comparisons (Lenth, 2016). We controlled for baseline pupil size in all analyses involving infants' pupil diameter. A standard value of $p = .05$ was used to define statistical significance. Finally, the pupillometry study and analysis plan were preregistered. However, as might be expected when working with a new methodology for the first time, there were a few minor deviations from the preregistration. None of these deviations influenced our general pattern of results, but nonetheless details of these deviations are described in further detail in the supplementary material.

Validity checks: did the data behave as expected?

Because the pupillometry technology used in this research is new, we first examined the data to explore the extent to which certain features followed expected patterns. We began by visually examining the data recorded by the Pi and processed in

Matlab to investigate the extent to which infants showed typical patterns of habituation. If the data revealed that infants' attention attenuated across time, with progressive reductions in looking as the six blocks of stimulus presentation proceeded, it would increase our confidence that the technology was appropriately capturing infants' looking behavior in relation to the presentation of stimuli. To perform this analysis, we calculated the proportion of time that Pi/Matlab detection of pupils indicated that infants spent looking during each trial that they viewed. Because we were primarily interested in infants' attention to the activity occurring in the video rather than the grey screen or still frame, we limited our analyses to only the video portion of each trial (although we note that we observed the same pattern of results if we include infants' looking to the grey screen and still frame). First, any frame for which the Matlab program detected a pupil was classified as "looking" and any frame for which Matlab did not detect a pupil was classified as "not looking." We then created a "proportion of time spent looking" measure by summing the number of frames participants spent looking during a given trial and dividing that value by the total number of frames in the trial. Infants' interest in the videos was high overall – on average, infants spent 93% (*SD* = 11%) of each trial looking to the screen. However, the proportion of time spent looking to each trial did decrease across the six blocks. In a linear mixed-effects model with a fixed effect of block and random intercept for subjects, we found a significant linear trend, $\beta = -0.04$, $t(679.74) = -4.05$, $p < .001$. As can be seen in Figure 3.3, the proportion of time infants spent looking to the video decreased across the six blocks (with some recovery in looking at the sixth block). This result provided one source of validation that the pupillometry methodology was indeed appropriately detecting infants' pupils.

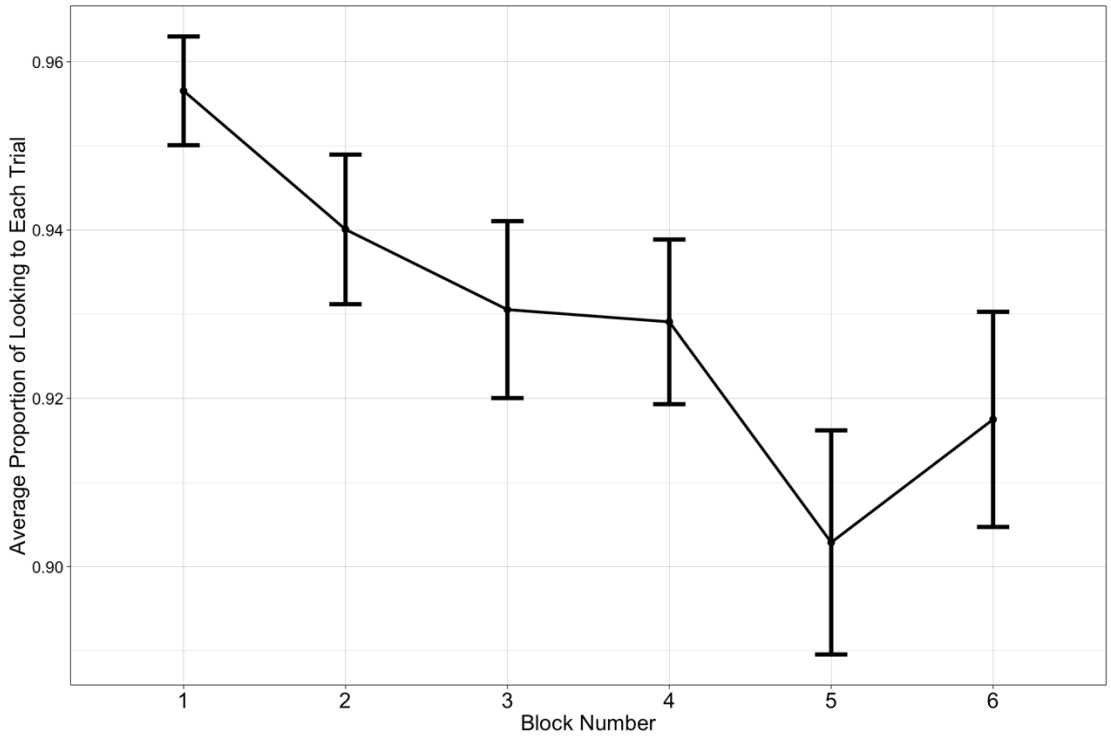


Figure 3.3. Average Pi-recorded proportion of time infants spent looking to the stimulus video across the six blocks. Error bars indicate +/- 1 SE.

We next examined infants' PDR to salient moments in the activity stream. Because infants' pupil size changes in response to cognitive stimuli (see Hepach & Westermann, 2016 for a review), we expected infants' pupils to dilate in response to the appearance of the still frame after the grey screen as well as to the start of the video after the still frame. As in the previous analysis, we began by visually examining infants' average pupil size across the grey screen, still frame, and video onset (see Figure 3.4). Upon visual examination, we noted three primary features: (1) it appears that there is a pupillary light reflex (PLR) in response to the still following the grey screen, (2) the patterns of pupil dilation in response to the video onset suggest sensitivity to the change in cognitive stimuli, and (3) we see a pattern of anticipation for both still frame and video

onset during the preceding time window (e.g., increasing pupil size at the end of the grey screen and still frame).

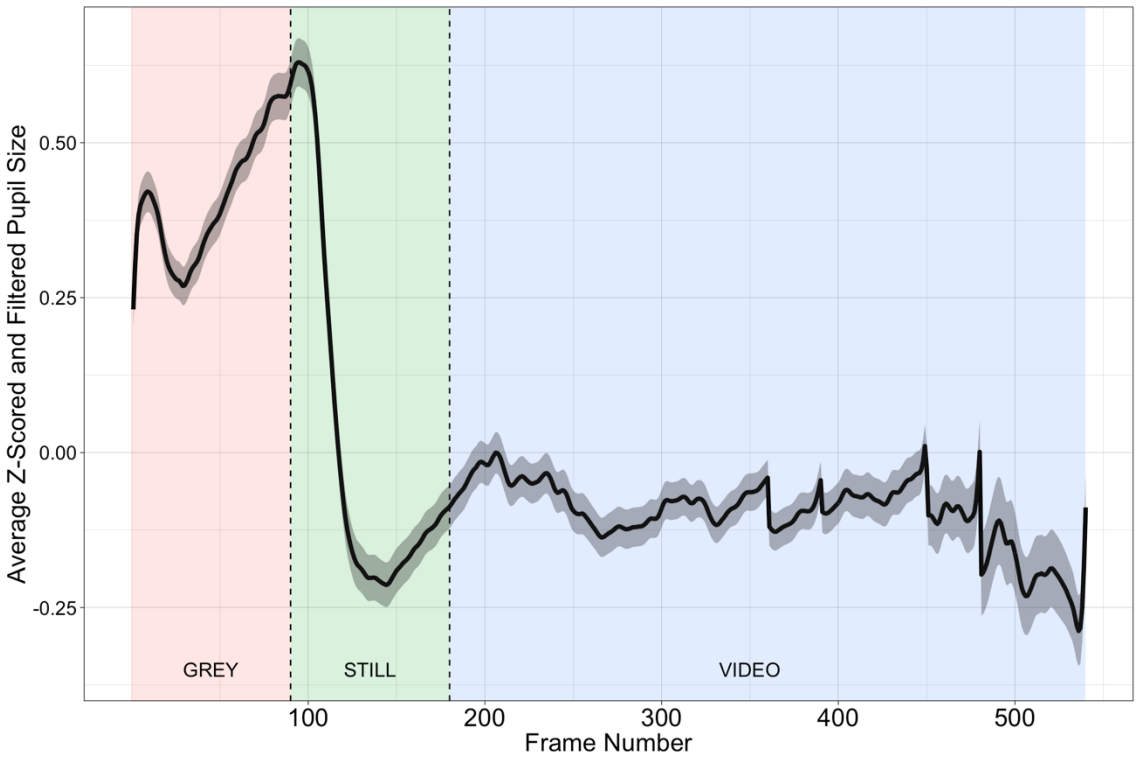


Figure 3.4. Average z-scored, filtered pupil size across the course of the videos. This figure depicts all videos averaged together. The dashed vertical lines represent the offset of the grey screen and the offset of the still frame, respectively. The grey screen, still frame, and video regions are indicated by background color. Shading around lines indicates +/- 1 SE.

To explicitly test for a PDR to the onset of the still frame, we defined two areas of interest: one second before the still frame (i.e., while the screen was still grey) and one second after appearance of the still frame. If infants were responding to the onset of the still frame, we would expect their pupil diameter to be smaller to the one second region prior to still-frame onset than to the one second period after onset of the still frame. This linear mixed-effects model included a fixed effect of video region (pre- vs. post-still

onset), random intercepts for subjects and videos. Contrary to our original predictions but consistent with a PLR, infants' pupil diameter was larger in the one second prior to the still-frame onset (M = 0.53, SD = 0.96) than in the one second region after the still-frame onset (M = 0.37, SD = 1.06), β = -0.08, $t(43,114) = -18.17$, $p < .001$. As described previously, the change from the grey screen to a still of the first frame in the video (which depicted a woman wearing a light blue shirt and seated in front of a black screen) evoked a very brief increase, then rapid constriction of the pupil. This pattern of constriction immediately after the still-frame onset likely reflected a pupillary light reflex (Laeng, et al., 2012; Loewenfeld, 1993; Binda, Pereverzeva, & Murray, 2013) – rapid, large pupil constriction in response to a change in luminance – rather than a response to cognitive stimuli. To directly test the influence of luminance, we ran a linear mixed-effects model predicting pupil size from a fixed effect of z-scored luminance, random intercepts and slopes for subjects, and random intercepts for videos. Luminance was significantly, negatively predictive of pupil size, β = -0.14, $t(25.85) = -5.09$, $p < .001$. Consistent with the PLR, as luminance increased from grey screen to still frame, pupil size correspondingly decreased.

Our next analysis explored observers' PDR as an index of a cognitive response to stimulus onset (or the juncture at which the actor began to move after the still frame). As can be seen in Figure 3.4, there did *not* appear to be a PLR in response to luminance changes at the start of the video. This was expected as the general scene did not change; the actor simply began to move. To explore the effect of this movement, we ran a linear-mixed effects model predicting pupil size from a fixed effect of video region (pre or post video onset) and random intercepts for subjects and videos (including a random slope for

subjects, as we did in the previous model, caused issues with model convergence). We found that infants' pupil size was larger post-video onset (M = -0.03, SD = 0.90) than pre-video onset (M = -0.14, SD = 0.92), β = 0.05, t(43,114.17) = 12.40, p < .001. As predicted, infants' pupil diameter increased with a change in cognitive stimuli (i.e., when the actor began to move). However, as can be seen in Figure 3.4, pupil size began increasing even before the start of the video. This is likely due to both recovery from the pupillary light response and infants' ability to anticipate movement beginning as they gained experience of the pattern of stimulus presentation (e.g., grey screen followed by still frame followed by video) over the course of the study.

In our final data checking procedure, we hand coded seven (25%) of the video files and explored the extent to which (1) there was agreement in frame-by-frame looking and looking away across the Pi-detected and Matlab-coded versus hand-coded video files, and (2) whether deleting "false alarm" trials (trials on which the Pi and Matlab programs detected that the infant was looking while the hand coder indicated that the infant was looking away) influenced the general pattern of observed results. To assess agreement across Pi/Matlab-coded and hand-coded videos, we calculated the number of hits (Pi/Matlab indicated infant *looking* and hand coder also indicated *looking*), misses (Pi/Matlab indicated that the infant was *looking away* but the hand coder indicated that the infant was *looking*), false alarms (Pi/Matlab indicated that the infant was *looking* but the hand coder indicated that the infant was *looking away*), and correct rejections (Pi/Matlab indicated that the infant was *looking away* and the hand coder also indicated that the infant was *looking away*). These numbers aggregated across the seven files are reported in Table 3.2; by-infant proportions are available in the supplementary materials.

We found strong agreement between the Pi/Matlab-coded and hand-coded videos, with looking judgments corresponding (i.e., hits + correct rejections) on 94.8% of frames. Disagreements were relatively rare, with false alarms occurring more frequently than misses (3.3% vs. 1.9% of the data, respectively).

Table 3.2.

Total number of hits, misses, false alarms, and correct rejections across the seven Pi/Matlab coded and hand-coded videos. The percentages represent the proportion of the total (N = 58,730) hand-coded frames for which each type of match or mismatch occurred.

	Hand: Looking	Hand: Looking Away
Pi: Looking	52,148 (Hits: 88.8%)	1,953 (False Alarms: 3.3%)
Pi: Looking Away	1,096 (Misses: 1.9%)	3,533 (Correct Rejections: 6%)

In a separate set of analyses, we analyzed the data in the manner described below for only the seven participants for whom we had information about the accuracy between the Pi/Matlab-coded and hand-coded data. The results of these analyses generally paralleled those reported below, even with only 25% of the data, suggesting robust findings. We then explored the consequence of mismatches between the Pi/Matlab-coded and hand-coded data. Specifically, we removed any "false alarm" frames[6] – frames for which the hand coder indicated that the infant was not looking while the Pi/Matlab program found a pupil. We then re-analyzed the data and compared these findings to the results without false alarms removed. The results were strikingly similar, both in the results of significance tests and in visual comparison of plots of the findings. More information regarding these comparisons is available in the supplementary materials.

[6] Because "misses" are frames for which the Matlab program did not detect a pupil (and thus there was no estimate of raw pupil diameter), it was not possible to analyze the data with and without "misses."

In sum, the preceding analyses provided clear validation of Pi/Matlab-based coding of infant looking behavior. Thus, these data were used in all subsequently reported analyses.

Did motionese enhance infants' overall attention to action?

We explored the extent to which motionese, relative to adult-directed action, influenced infants' overall attention to unfolding activity by examining both (1) infants' looking duration to motionese versus adult-directed action, and (2) their overall average pupil diameter (i.e., tonic pupil size) in response to motionese versus adult-directed action. For these analyses we focused only on frames corresponding to the video portion of each trial (i.e., ignoring the grey screen and still frames) as this was where the difference between motionese and adult-directed action should emerge most clearly given the stimuli employed in the present study.

For analysis of infants' looking to motionese versus adult-directed action, any frame for which the Pi/Matlab program detected a pupil was classified as "looking" and any frame for which Pi/Matlab did not detect a pupil was classified as "not looking." We then created a proportion of time spent looking to each trial by summing the number of frames participants spent looking during each trial and dividing that value by the total number of frames in the trial. We conducted a linear mixed-effects model predicting the proportion of time spent looking to a given trial from a fixed effect of demonstration type (i.e., motionese versus adult-directed action) and random intercepts for subjects and videos. Although means were in the predicted direction, contrary to our prediction we found no significant difference in the proportion of trials for which infants looked to

motionese (M = 94%, SD = 10%) versus adult-directed action (M = 93%, SD = 12%), β = -0.007, $t(10.81) = -1.33$, $p = .21$. Because infants were looking for nearly the full duration of each trial, it is possible that a ceiling effect prevented detection of the predicted difference in looking to infant- versus adult-directed action[7].

We previously demonstrated that infants' looking declined over the course of the six blocks, and we hypothesized that a difference in infants' looking to motionese relative to adult-directed action might have emerged for later blocks for which the ceiling effect might not have been as pronounced. Visual inspection of Figure 3.5 is consistent with this prediction. Infants' pupil size in response to motionese is higher than to adult-directed action in five out of the six blocks, and pupil size in response to adult-directed action is never higher than pupil size in response to motionese (i.e., for the one block in which pupil size is not larger to motionese pupil size is approximately equal across conditions). Further, it appears that the magnitude of this difference was larger for later blocks, when infants were looking less overall, perhaps because the ceiling effect began to decrease.

We thus conducted this analysis again including an effect of block and an interaction between block and demonstration type. In a model controlling for block and the interaction between block and demonstration type, the p-value associated with the main effect of demonstration type decreased, β = -0.007, $t(660.82) = -1.87$, $p = .06$. This model also replicated our earlier analysis of looking time decreasing over the six blocks – we found a significant linear trend for block, β = -0.04, $t(683.82) = -4.01$, $p < .001$.

[7] A possible concern is that our method of removing participants who were not attending for 50% of a given trial influenced this comparison. These mean looking to infant- and adult-directed action decreased only slightly (to 89.4% and 88%, respectively) and the general pattern of results holds when we include these "low-looking" participants in the analysis. Thus, our method of excluding "low-looking" participants does not seem to explain the observation of high looking overall.

However, the interaction between demonstration type and block was not significant, $F(5, 660.31) = 0.35$, $p = .88$.

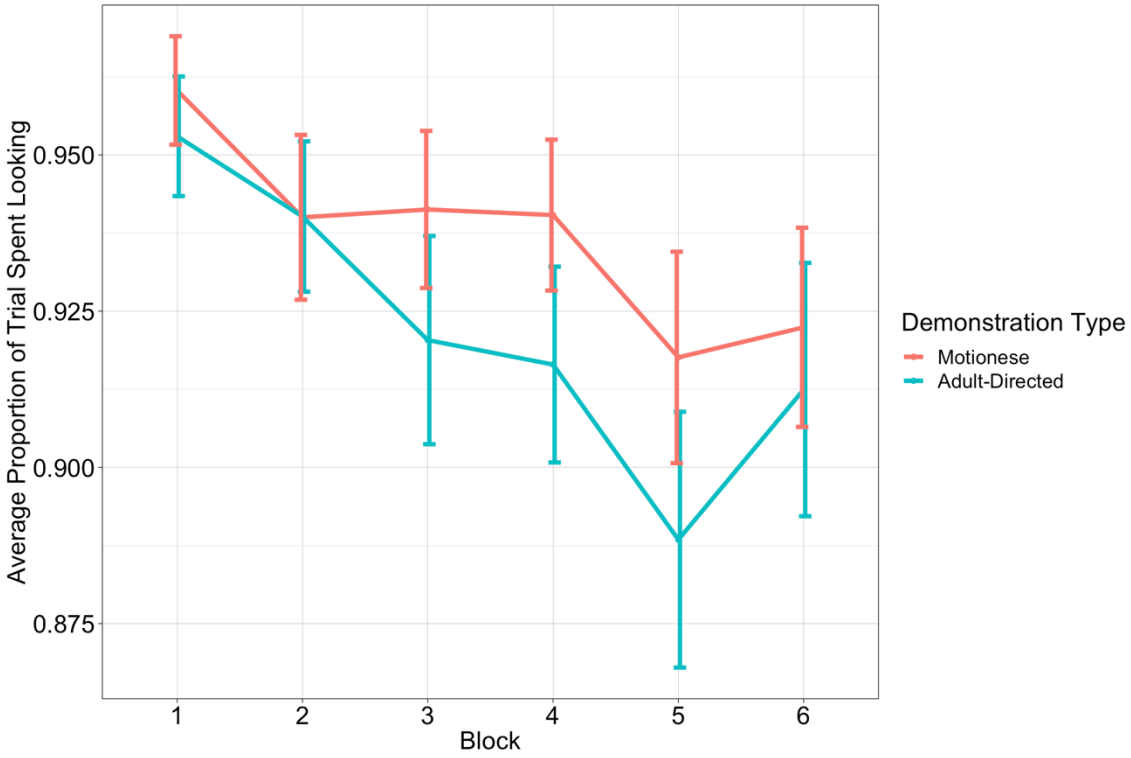


Figure 3.5. Average proportion of time infants spent looking to the stimulus video across the six blocks and by demonstration type. Error bars indicate +/- 1 SE.

Next, we turned to examining the influence of motionese versus adult-directed action on infants' tonic pupil size. For this analysis, we ran a linear mixed effects model predicting infants' z-scored, filtered pupil diameter from a fixed effect of demonstration type (motionese versus adult-directed action) with random effects of subjects and videos. We additionally controlled for infants' baseline pupil diameter. This analysis is depicted in Figure 3.6. Again, contrary to our predictions, we did not find a significant effect of demonstration type, $\beta = -0.04$, $t(9.61) = -1.38$, $p = .20$, though infants' average pupil

diameter tended to be larger in response to motionese (M = -0.02, SD = 0.80) over adult-directed (M = -0.15, SD = 0.81) activity sequences.

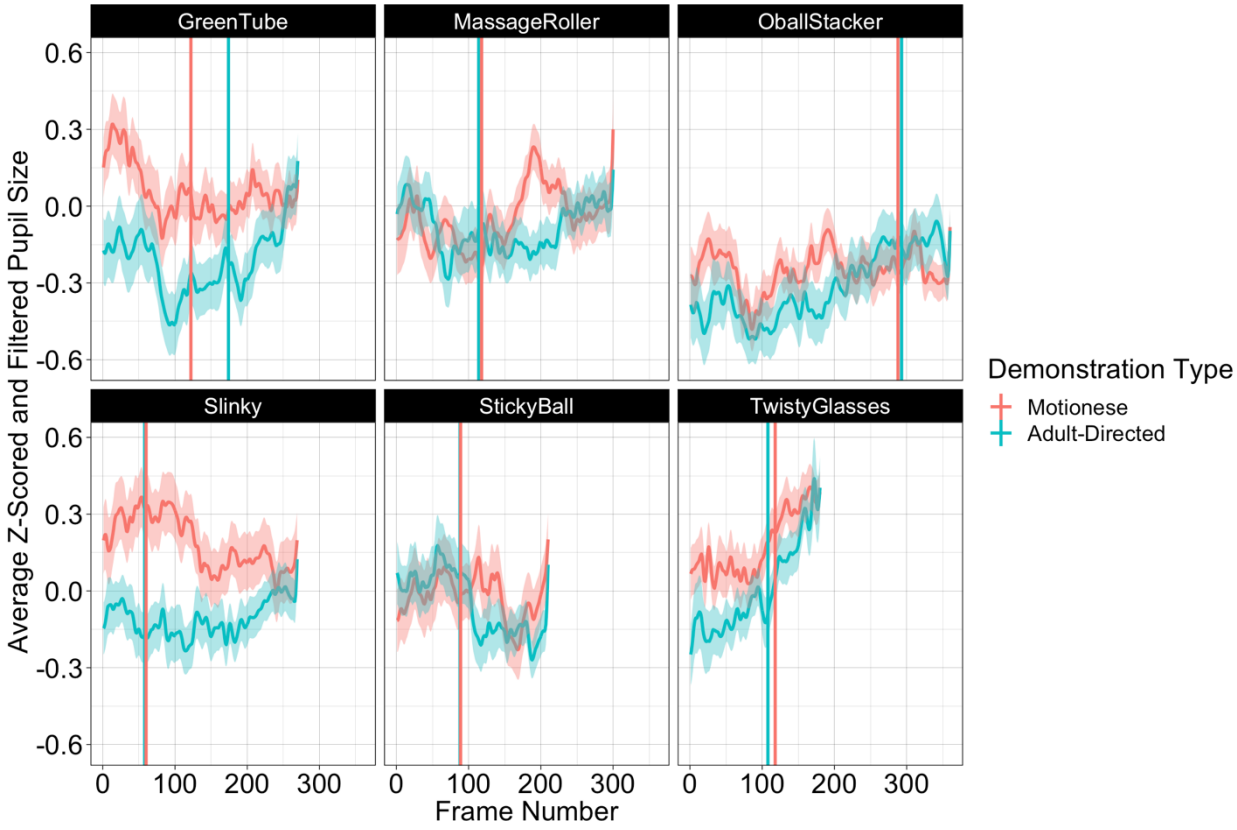


Figure 3.6. Average z-scored, filtered pupil size to motionese and adult-directed versions of each video. Vertical lines represent the location of the one major action boundary, and color of the line indicates whether the boundary occurred in the motionese or adult-directed version of the object demonstration. Shading around lines indicates +/- 1 SE.

Did infants selectively attend to action boundaries in continuous activity sequences?

In our next set of analyses, we explored the extent to which infants preferentially attended to boundaries in unfolding activity sequences, as indexed by changes in pupil diameter. For these analyses, we focused in particular on activity surrounding the one major action boundary depicted within each video. As described in Chapter II, we defined pre-boundary, boundary, and post-boundary regions in each video. The pre-

boundary region covered the one second of activity (or 30 frames) occurring prior to the action boundary. The boundary region began at the action boundary and extended for the next one second (30 frames), and the post-boundary region began at the end of boundary region and continued 1 additional second, or 30 more frames. In previous research exploring adults' PDR to action boundaries (e.g., Tanaka and colleagues, in preparation), researchers used half-second pre-boundary, boundary, and post-boundary regions. However, this time window might miss infants' pupillary response to the action boundary, because there is evidence that infants' pupils respond to cognitive stimuli more slowly than adults' (e.g., Verschoor, Spapé, Biro, & Hommel, 2013; Verschoor, Paulus, Spapé, & Hommel, 2015; Zhang, Jaffe-Dax, Wilson, & Emberson, 2018). Thus, we opted to extend the windows to one-second regions. This timing is also consistent with prior work in which researchers incidentally provided information about the timing of infants' response to action boundaries. Jackson and Sirois (2009) measured infants' PDR to a train entering and emerging from a tunnel. Visual examination of infants' PDR to the boundary at which the train emerged from the tunnel suggested that the response peaked and began returning to baseline within one second after the action boundary. Infants' response to the onset of the video in our stimuli also seems generally consistent with this one-second window. As can be seen in Figure 3.4, the peak and start of the return to baseline in infants' PDR to the video onset occurs approximately one second after the start of the video, further supporting the inference that the timing of infants' response to cognitive stimuli is best examined in a one-second window after a cognitive event (e.g., a video onset or an action boundary).

To test for a possible boundary-related PDR across all videos (regardless of whether activity depicted was motionese versus adult-directed action), we ran a linear mixed effects model predicting z-scored, filtered pupil diameter from a fixed effect of region (pre-boundary, boundary, post-boundary) and random intercepts for subjects and videos. We additionally controlled for baseline pupil size. Because we were specifically interested in the boundary region, the video frames included in these analyses were limited to those occurring in pre-boundary, boundary, and post-boundary regions. Video frames outside of these regions were eliminated from the current analyses. We found no significant main effect of region, $F(1, 62{,}600) = 2.50$, $p = .08$. On average, infants' pupil diameter did not differ significantly across pre-boundary, boundary, and post-boundary regions. This finding indicated that infants failed to display a systematic boundary-related PDR when considering their response to both motionese and adult-directed activity sequences taken together.

Did motionese enhance infants' response to boundaries within continuous activity?

To explore the extent to which motionese influenced infants' response to boundaries, we ran the same mixed-effects model described above, but now including fixed effects of demonstration type and an interaction between region and demonstration type, while still controlling for baseline pupil size. As in previous analyses, we found no significant main effect of demonstration type, $F(1, 10) = 1.18$, $p = .30$, or region, $F(2, 62{,}598) = 2.53$, $p = .08$. However, we did observe a significant interaction between demonstration type and region, $F(2, 62{,}598) = 11.11$, $p < .001$. To explore this interaction, depicted in Figure 3.7, we ran two separate mixed-effects models for

motionese and adult-directed demonstrations. In adult-directed demonstrations, there was no effect of region, $F(2, 31{,}376) = 1.89$, $p = .15$. In contrast, for motionese demonstrations we observed a significant effect of region, $F(2, 31{,}196) = 13.43$, $p < .001$, that followed both significant linear, $\beta = 0.03$, $t(31{,}195.80) = 4.15$, $p < .001$, and quadratic trends, $\beta = -0.02$, $t(31{,}195.80) = -3.10$, $p = .002$. PDR to pre-boundary slides ($M = 0.01$, $SD = 0.81$) was lower than PDR to both boundary slides ($M = 0.05$, $SD = 0.79$), $\beta = -0.04$, $p < .001$, and post-boundary slides ($M = 0.04$, $SD = 0.76$), $\beta = -0.04$, $p < .001$. However, PDR did not differ significantly between boundary and post-boundary slides, $\beta = 0.01$, $p > .99$. To summarize, in response to motionese demonstrations, infants' pupil size increased within boundary regions (relative to pre-boundary regions) and remained high post-boundary. These effects were not observed in infants' PDR to adult-directed demonstrations.

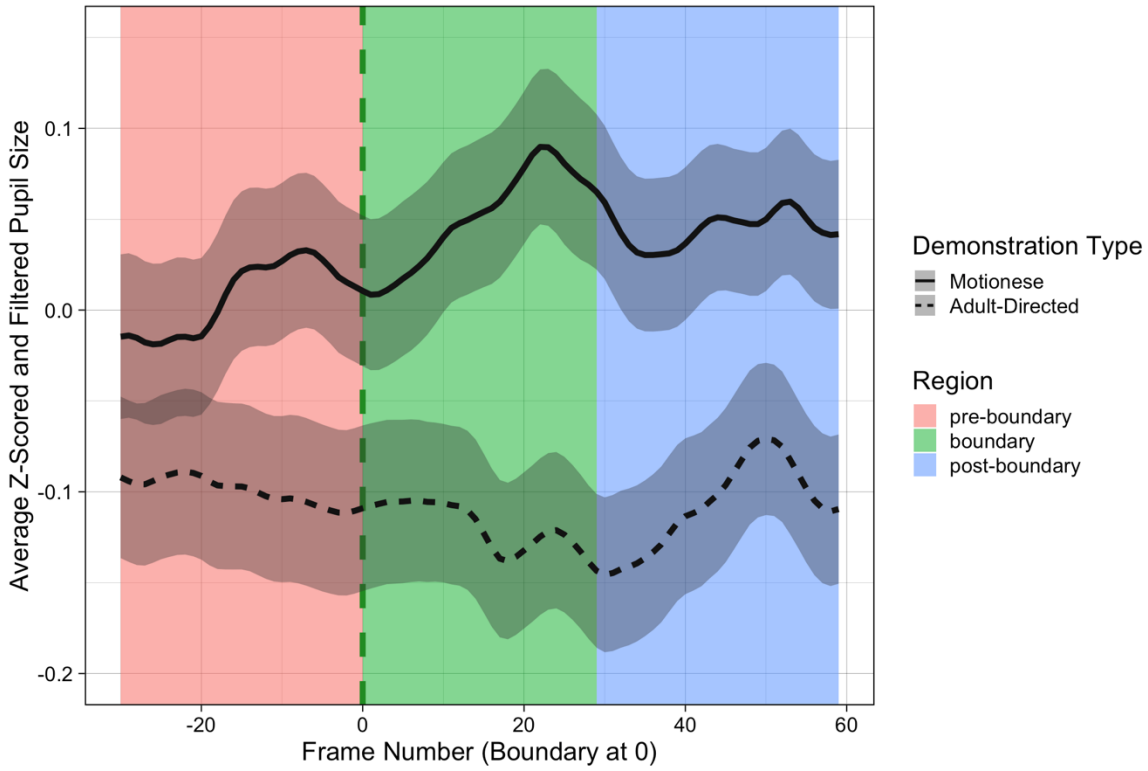


Figure 3.7. Average z-scored, filtered pupil size in response to motionese (solid line) and adult-directed (dashed line) action. Video region is indicated by the background color of the plot, with the boundary occurring at time 0 on the x-axis. Shading around lines indicates +/- 1 SE. This pattern plotted separately for each infant is available in the supplementary materials.

In previous research, we've found that adults attention to the structure of unfolding, novel action emerges across repeated viewing (e.g., Kosie & Baldwin, 2019a). We next conducted an exploratory analysis to investigate infants' processing of structure across repeated viewing in another linear mixed-effects model for the motionese action condition only (because the boundary effect was not significant in adult-directed condition). In this model we included fixed effects of region (pre-boundary, boundary, post-boundary) and block, random intercepts for subjects and videos, and controlled for baseline. We found that the main effect of region was no longer significant when including block and its interaction with region, $F(2, 31{,}193) = 0.82$, $p = .44$. However, we did find a significant effect of block, $\beta = -0.03$, $t(31{,}204.30) = -13.00$, $p < .001$, suggesting that the magnitude of infants' PDR differs across the six blocks. This is consistent with previous work suggesting that infants' (and adults') PDR habituates over time (e.g., Bala et al., 2016). We additionally found a significant interaction between region and block at the quadratic level, $\beta = -0.007$, $t(31{,}192.79) = -1.93$, $p = .05$. To follow up on this interaction, we ran separate linear mixed-effects models for each block in which we predicted infants' pupil size from fixed effects for region and random effects for subjects and video, again controlling for baseline pupil size. To control for multiple comparisons, all p-values were Bonferroni corrected. Although visual examination of pupil size patterns (depicted in Figure 3.8) suggested that there was pupil dilation during boundary regions as early as the first or second block, a *significant* effect of boundary

region on pupil dilation patterns did not emerge until the fourth block. In block four, we found both significant linear, $\beta = 0.05$, $t(5{,}012.97) = 3.46$, $p = .006$, and quadratic trends, $\beta = -0.05$, $t(5{,}012.97) = -3.52$, $p = .005$. Pupil size was smaller pre-boundary than at both the boundary and post-boundary regions, $ps < .002$. However, pupil size did not differ between the boundary and post-boundary regions, $p = .56$, suggesting that infants' pupil dilation increased at the boundary and remained high post-boundary. In block five, we found only a significant quadratic trend, $\beta = -0.05$, $t(4{,}476.88) = -3.25$, $p = .01$. Pupil size was significantly lower pre-boundary than at the boundary region, $p < .001$. However, pupil size did not differ between the pre-boundary and post-boundary regions nor between boundary and post-boundary regions, $ps > .13$. Again, infants' pupil size increased at the boundary and remained high post-boundary. The non-significant pre- to post-boundary comparison suggested that, on block five, infants' post-boundary pupil size did decrease (making it closer to pre-boundary pupil size) but not enough to reveal a significant difference in boundary and post-boundary pupil size. We found no significant region effects in the sixth block, $ps > .74$, but this may have been due to the inclusion of fewer infants in this final block and thus lower power.

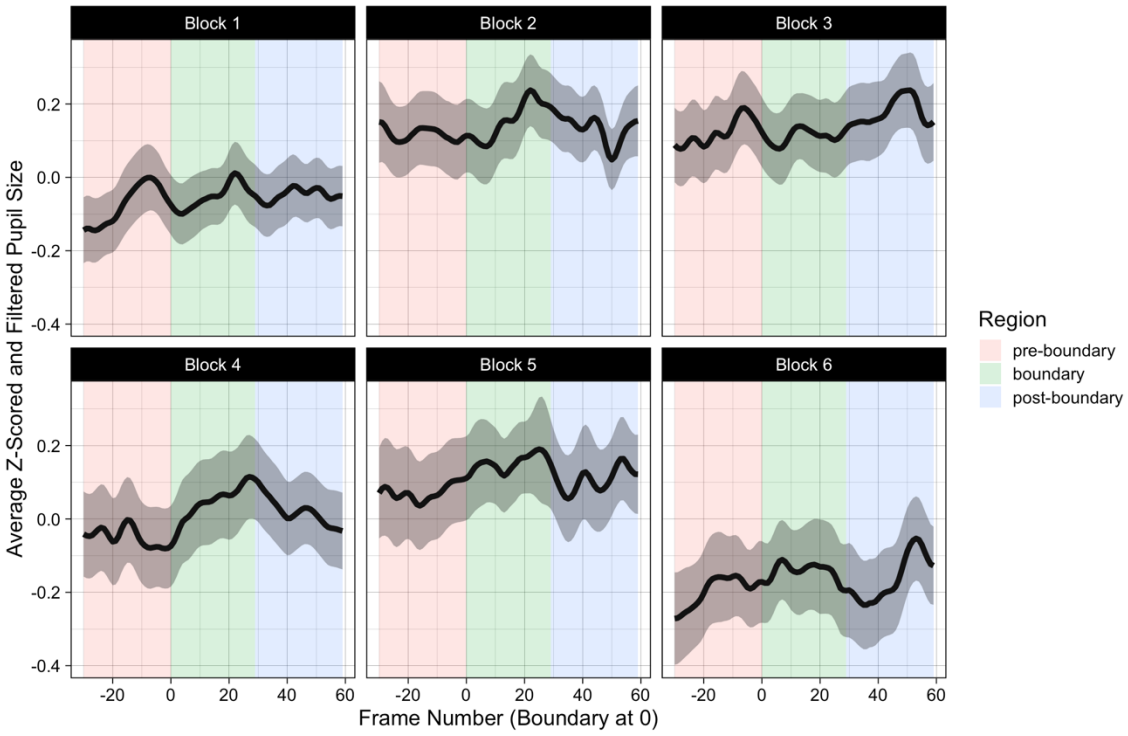


Figure 3.8. Average z-scored, filtered pupil size in response to motionese action across the six blocks. Video region is indicated by the background color of the plot, with the boundary occurring at time 0 on the x-axis. Shading around lines indicates +/- 1 SE.

Did luminance predict PDR above and beyond effects of demonstration type and region?

Recall that when we analyzed luminance in Chapter II, we observed a significant interaction between demonstration type and video region; luminance was lower at boundary than pre- and post-boundary regions, but in motionese demonstrations only. There could be reason for concern that these luminance patterns seem to be in line with our observed PDR results. Specifically, it is known that pupil diameter decreases as luminance increases and, correspondingly, increases as luminance decreases (e.g., Loewenfeld, 1993). Therefore, our observation of larger pupil size within the boundary regions of motionese action might be related to lower luminance in that region. To control for this potential confound, we ran the demonstration type by video region

analysis again with the same fixed and random effects structure and controlling for baseline, but this time also controlling for z-scored luminance of the video frames. Consistent with our model that did not control for luminance, we observed a significant interaction between demonstration type and region, $F(2, 62{,}600) = 11.25$, $p < .001$, but no main effect of demonstration type or region independently, $ps > .08$. Additionally, in this model, luminance alone was not predictive of pupil size, $\beta = -0.004$, $t(58{,}868.78) = -1.20$, $p = .23$. Despite the correspondence in luminance and PDR patterns, we found that luminance was not predictive of pupil size above and beyond effects of demonstration type and video region, thus our PDR results held even when controlling for video luminance. Pre-boundary, boundary, and post-boundary luminance and PDR patterns across motionese and adult-directed action are plotted in Figure 3.9.

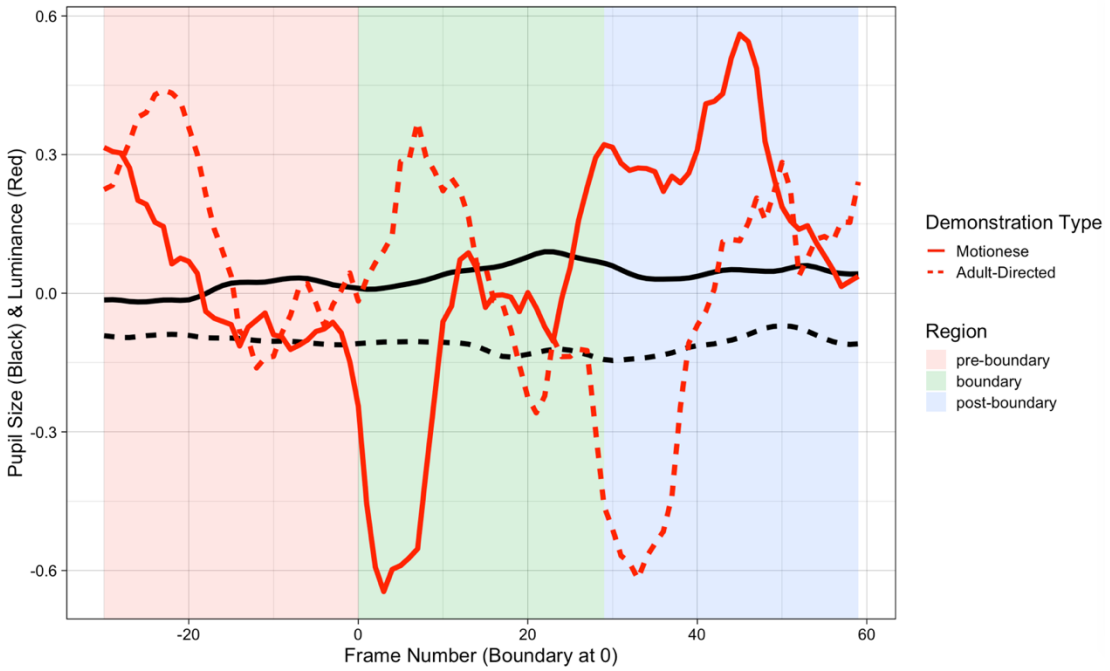


Figure 3.9. Z-scored and filtered pupil size (plotted in black) and z-scored video luminance (plotted in red) to pre-boundary, boundary, and post-boundary regions of motionese (solid line) and adult-directed (dashed line) action. These figures are plotted without a measure of error to facilitate interpretability. However, these effects with error are plotted separately above (pupil size in Figure 3. 7 and luminance in Figure 2.4).

Was infant age predictive of looking time and PDR patterns above and beyond effects of demonstration type and video region?

In our final analysis of control variables, we asked whether infant age might have influenced any of the observed effects. We first re-ran the analysis predicting infants' looking to motionese versus adult-directed demonstrations. This model had the same fixed and random effects structure as the previous model, with an additional fixed effect of infant age (in days). Again, we observed only a linear trend for block, $\beta = -0.04$, $t(652.78) = -4.12$, $p < .001$, suggesting that looking decreased throughout the session. No other effects, including infant age, were significant, $ps > .17$. Next, we re-ran the analysis examining the effects of demonstration type, video region, and their interaction. We again used the fixed and random effects structure described earlier, with the addition of a fixed effect of infant age. Infant age was not predictive of pupil size, $\beta = -0.0001$, $t(24.51) = -0.10$, $p = .92$. As in previous analyses, the only significant effect in this model was the interaction between demonstration type and video region, $F(2, 62{,}600) = 11.25$, $p < .001$. Taken together, it appeared that infant age was not predictive of pupil size in any of the analyses, nor did controlling for age impact the observed results.

Did infants interact with objects more when they had previously viewed them in motionese demonstrations?

Our final set of analyses focused on infants' opportunity to play with the objects they had viewed in the video stimuli. Recall that each infant saw videos featuring all six objects, but that the identity of the interaction partner in these videos differed across infants. Therefore, a given infant saw three objects in motionese demonstrations and three objects in adult-directed demonstrations. They then had the opportunity to interact with

the six objects. The objects were presented in pairs that included one object that had been featured in a motionese demonstration and one that had been featured in an adult-directed demonstration. As in Chapter II, we examined infants' interest in these objects in four different ways. For a given pair of objects we coded the first object infants looked to after being presented with the pair, the length of time for which infants looked to each object in the first three seconds (the "looking-alone" phase), the length of time for which infants interacted with the object during the next twenty seconds (the "interacting" phase), and we also collected subjective judgments from coders regarding which object in each pair infants seemed to prefer. We compared these observations for objects that had been featured in motionese versus adult-directed demonstrations.

First, we examined the proportion of times a given object was the target of infants' first look. In contrast to the results presented in Chapter II, a chi-square test revealed no significant differences in the identity of objects that infants looked to first, $\chi^2(5) = 5.2$, $p = .39$ (see Figure 3.10). Consistent with our observations in Chapter II, the Oball™ Stacker was the most highly preferred object, with infants looking to it first on 67% of trials in which it was presented. The least preferred toy was the Slinky, with infants looking to it on only 30% of the trials in which it was presented. In Chapter II the object least preferred was the Green Tube (in the current data, the Green Tube was preferred on about 52% of the trials in which it was presented). We additionally explored the extent to which infants' first look was to the object they viewed in motionese versus adult-directed demonstrations. On average, infants looked first to the object they had viewed in motionese demonstrations on 48% of trials and to the object they had viewed in adult-directed demonstrations on 52% of trials. In a one-sample t-test, we asked

whether the proportion of times infants looked to the object they had viewed in the motionese demonstrations differed from chance (50%). We found that the proportion of demonstrations in which infants first looked to the object that had been featured in the motionese demonstration did not significantly differ from chance, $t(26) = -0.36$, $p = .72$.

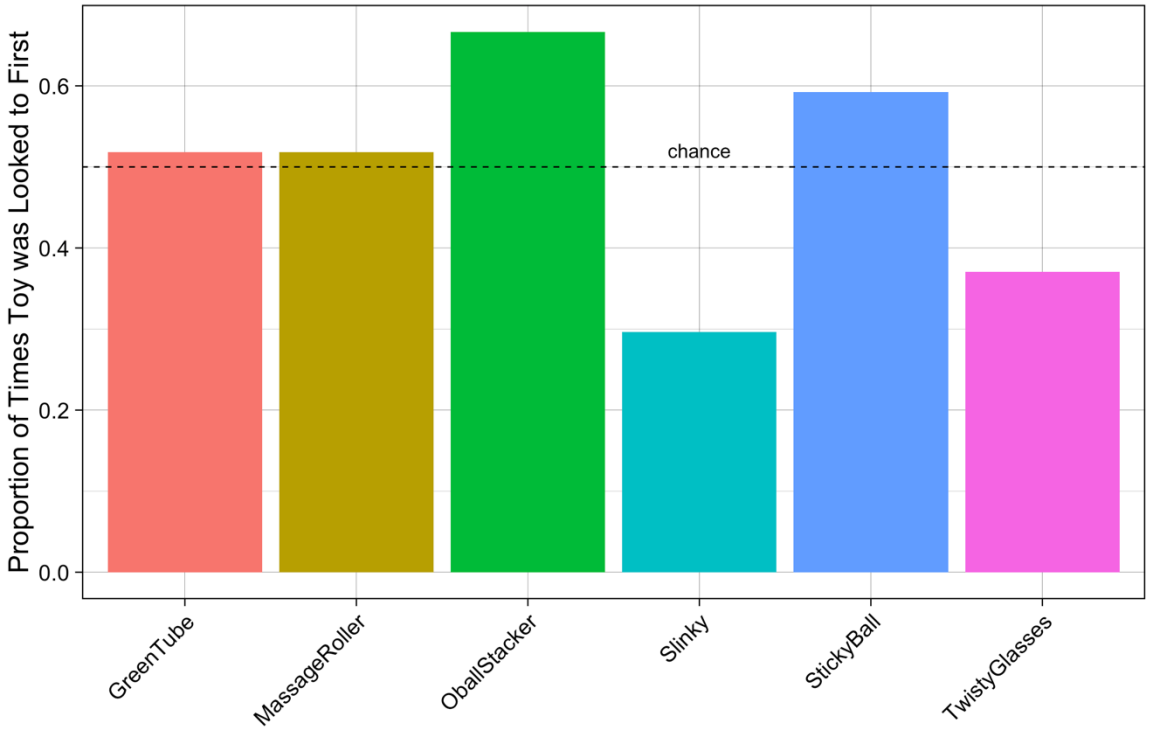


Figure 3.10. Proportion of trials (in which a given toy was presented) on which infants first looked to each object. The dashed line represents chance, which is .5 for any given object. As in Chapter II, figures depicting each object paired with each other object are available in supplementary materials.

Once the tray had been presented to infants, it was held just out of their reach for three seconds, and the duration of infants' looking to each object during that time period was coded. As in Chapter II, to examine overall differences in the amount of time infants spent looking to each toy, we ran a linear mixed-effects model predicting the number of seconds looking at a toy from a fixed effect of toy identity and random intercepts for subjects and object pair (whether it was the first, second, or third pair presented to the

infant). As can be seen in Figure 3.11, we again found significant differences in the amount of time infants spent looking across the six objects, $F(5, 133) = 2.69$, $p = .02$.

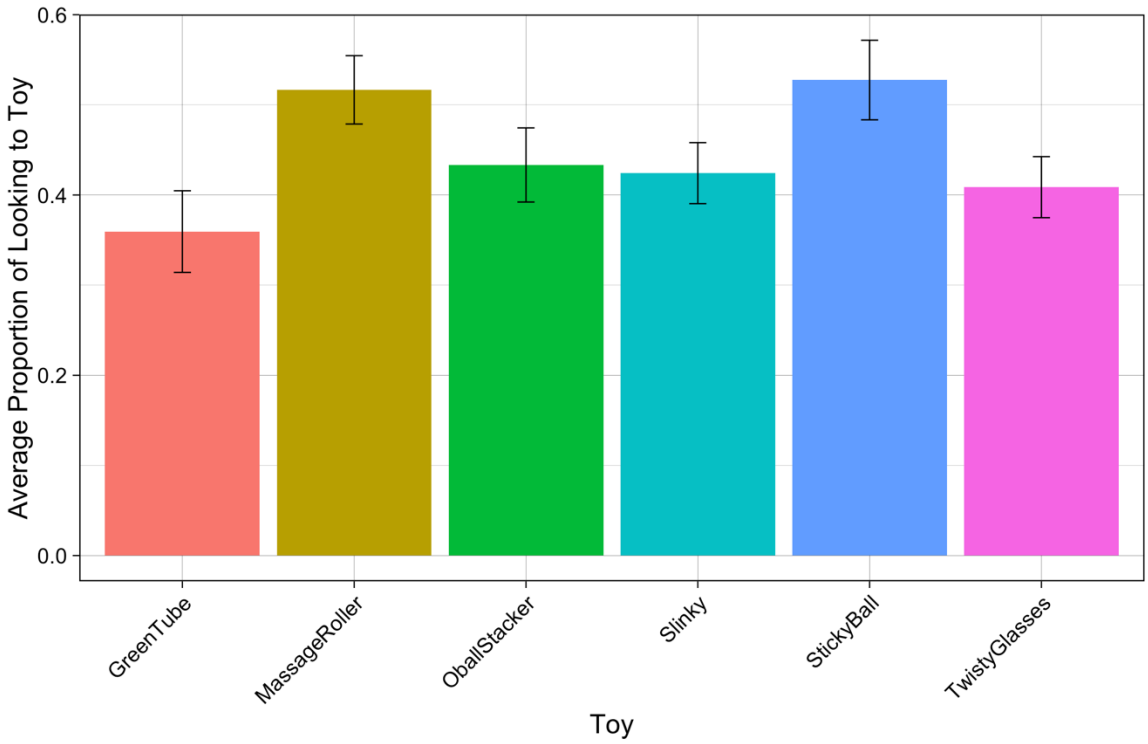


Figure 3.11. Proportion of three-second "looking" phase during which infants looked to each of the six objects. Error bars indicate +/- 1 SE. As in Chapter II, figures depicting each object paired with each other object are available in supplementary materials.

Infants were more likely to look at the Massage Roller and Sticky Ball than the other objects. However, there does not seem to be one object that is clearly preferred or ignored in this three-second "looking-alone" phase. We next explored the extent to which the proportion of time infants' spent looking to an object was dependent upon whether they had seen the object in a motionese or adult-directed demonstration. We ran a linear mixed effect model predicting the number of seconds infants spent looking to each object from a fixed effect of demonstration type (whether the infant had viewed that object in a

motionese or adult-directed demonstration) and a random intercept for subjects. Contrary to our predictions, infants spent more time looking to objects that they had seen in adult-directed (M = 1.46s, SD 0.64s) over motionese demonstrations (M = 1.21s, SD = 0.51s), β = 0.24, $t(137) = 2.49$, $p = .01$.

In our next analyses we explored the duration for which infants interacted with each of the six objects in the twenty-second "interacting" phase. Again, to test for differences in the amount of time infants spent interacting with each object, we ran a linear mixed-effects model predicting the number of seconds spent interacting with a toy from the toy identity (a fixed effect) and a random intercept for subjects (as in Chapter II, including an additional random intercept for object set as we did in the previous analysis caused issues with model convergence, so it was omitted from this analysis). As depicted in Figure 3.12, there were significant differences in the duration for which infants interacted with each object, $F(5, 106.15) = 7.05$, $p < .001$. As in Chapter II, infants seem to be less interested in the Twisty Glasses. Also, consistent with the "looking-only" phase, infants appear to be more interested in the Massage Roller and Sticky Ball (though in the "interacting" phase, the Green Tube was frequently chosen as well). However, there again isn't one object that stands out as overwhelmingly preferred. We also again ran a linear mixed-effects model exploring the effect of demonstration type (motionese or adult-directed) on the duration of time infants spent interacting with objects in which we included a fixed effect of demonstration type and a random intercept for subjects. We found no significant difference in the time infants spent interacting with objects that they had viewed them in motionese (M = 9.32s, SD = 6.29s) or adult-directed (M = 10.5s, SD = 6.31s) demonstrations, β = 1.13, $t(139) = 1.07$, $p = .29$.

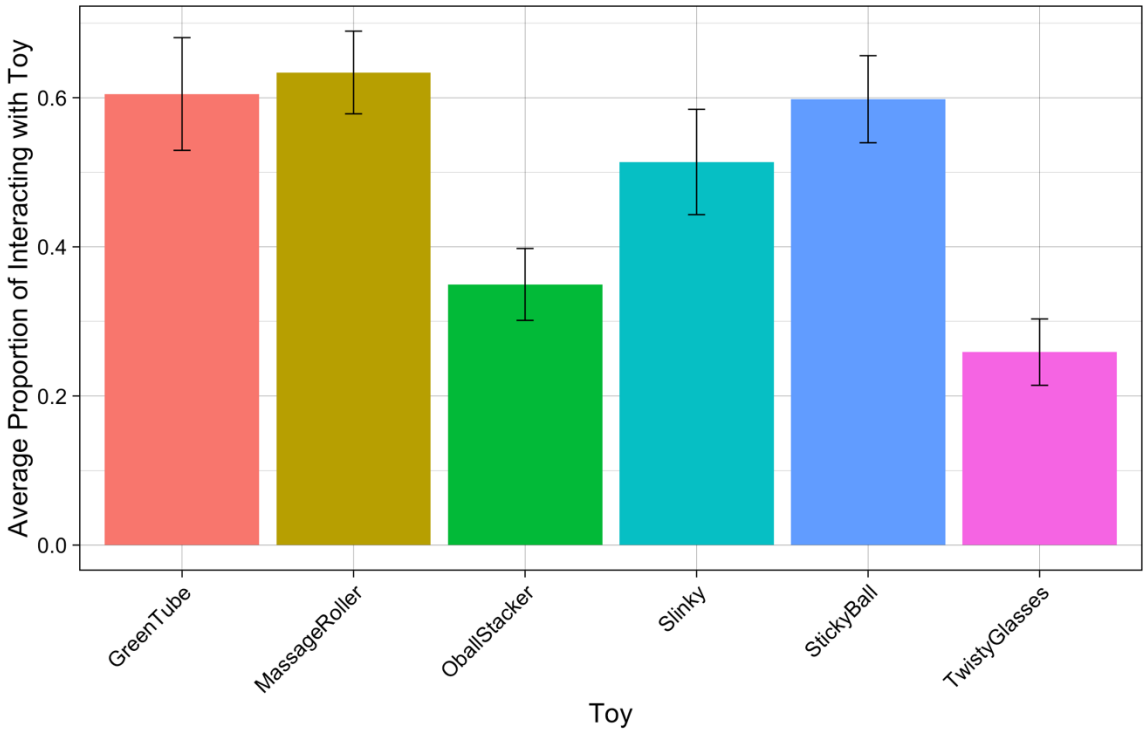


Figure 3.12. Proportion of twenty-second "interacting" phase during which infants were interested in each of the six objects. Error bars indicate +/- 1 SE. As in Chapter II, figures depicting each object paired with each other object are available in supplementary materials.

Finally, we explored coders' subjective judgments of the object that infants had preferred across the "looking-alone" and "interacting" phases. As depicted in Figure 3.13, there was again variability in the toys coded as "preferred" by infants. A chi-square test revealed significant differences in the identity of objects that were subjectively rated as being preferred by infants, $\chi^2(5) = 13.24$, $p = .02$. Consistent with prior analyses in this chapter, infants tended to prefer the Massage Roller and Sticky Ball and were less inclined to prefer the Twisty Glasses. As in previous analyses we also explored the extent to which the "preferred" object had been featured in a motionese or adult-directed demonstration. On average, infants "preferred" the object they had viewed in motionese demonstrations on 40% of trials and to the object they had viewed in adult-directed

demonstrations on 60% of trials. In a one-sample t-test, we found that the proportion of demonstrations in which infants preferred the object that had been featured in the motionese demonstration did not differ from chance, $t(26) = -1.78$, $p = .09$.

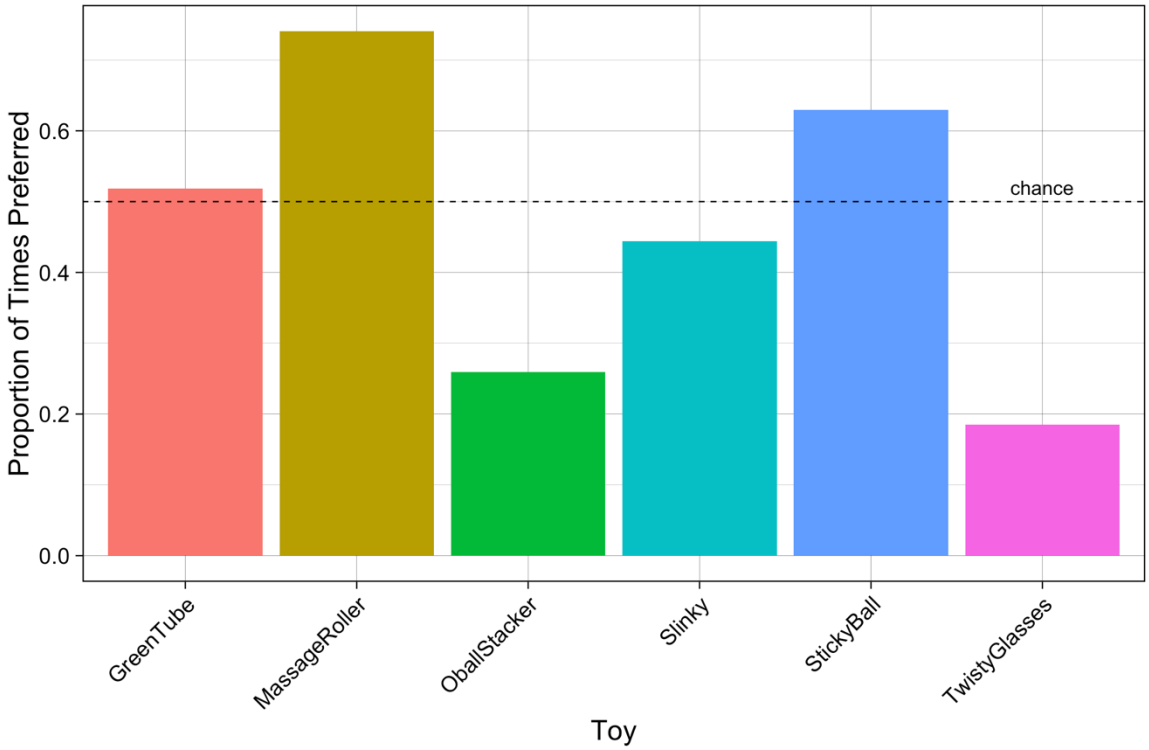


Figure 3.13. Proportion of trials (in which a given toy was presented) in which infants "preferred" each object. The dashed line represents chance, which is .5 for any given object. As in Chapter II, figures depicting each object paired with each other object are available in supplementary materials.

Discussion

To briefly review, we first performed a number of "validity checks" on the data to help validate the novel pupillometry methodology. We found expected patterns of reduction in infants' looking across trials. As well, PDR patterns generally conformed to predictions in response to the still frame and video onsets. Additionally, we found strong correspondence between Pi/Matlab-coded and hand-coded looking, suggesting that the Pi

camera and Matlab program accurately detected pupils during moments at which infants were looking at the screen.

We then explored the extent to which infants preferred to view motionese over adult-directed demonstrations, using both looking time and pupil size measures. Previous studies (e.g., Brand & Shallcross, 2008) found that infants prefer to view motionese over adult-directed demonstrations, and looking time means to motionese versus adult-directed videos in the present study trended this way, though the difference was not statistically significant. In particular, neither infants' looking, nor their pupil size were significantly greater in response to motionese demonstrations.

A subsequent set of analyses examined whether infants displayed a PDR to major action boundaries across the videos, as previously documented for adults by Tanaka and colleagues (in preparation). We measured infants' pupil size during pre-boundary, boundary, and post-boundary regions of unfolding activity sequences. Overall, infants' pupil diameter did not differ across pre-boundary, boundary, and post-boundary regions. However, when action was presented in a motionese format, infants' pupils displayed systematic increase in boundary relative to pre-boundary regions, and then remained high afterwards. Thus, infants indeed displayed a boundary-related PDR as was previously observed in adults, but for infants this was only the case for motionese demonstrations, supporting our prediction that motionese facilitates infants' detection of segmental structure within unfolding activity sequences.

Finally, we explored the influence of motionese on infants' interaction with novel objects. Overall, while we found some differences in infants' preference for objects, no single object stood out as being overwhelmingly preferred. These findings are consistent

with those reported in Chapter II. Additionally, during only the three-second "looking alone" phase, infants looked significantly longer at the object that had been demonstrated in the adult-directed demonstration. One possible explanation is that this was the result of a novelty preference – because infants were more interested in motionese videos, objects presented in adult-directed videos were more novel. However, this difference was not replicated in any of our other measures of infant interest, and we are thus cautious about interpreting this effect.

Regarding the "validity check," analyses overall increased confidence that the novel pupillometry methodology used in this research detected pupils appropriately and accurately measured infants' pupil diameter. One unexpected effect, the observation that pupil diameter was larger before the onset of the still frame, can likely be explained by luminance-related responding. Our original plan was to match the overall average luminance of the grey screen to the overall average luminance of the first frame in the video. Unfortunately, we subsequently learned that the luminance of the grey screen was altered due to video compression when combining the grey screen, still frame, and video files, and was not actually matched in luminance to the still frame in the stimuli viewed by infants (this issue is described in detail in the supplementary material). Although the PLR observed to the onset of the still frame after the grey screen was not predicted, it nevertheless was a sensible outcome given that the still frames displayed luminance increases relative to the grey screen, especially in certain regions of the video (e.g., the actor's face or shirt). Also sensible was the pattern of increases in pupil-diameter as the still images transitioned to the videos (i.e., when the motion began). Finally, analyses comparing Pi/Matlab-coded versus hand-coded looking behavior increased confidence in

the results reported here, as well as providing important validation of this relatively new pupillometry technology used with streaming visual stimuli.

A collection of analyses examined our first major research question: whether infants would prefer motionese over adult-directed action. One perspective on the outcome of these analyses is that we failed to replicate this previously observed preference (e.g., Brand & Shallcross, 2008), in that we found no difference in infants' looking duration nor in their average pupil size in response to infant- versus adult-directed demonstrations. Another, perhaps more nuanced, perspective on our pattern of findings is that we observed several hints of a motionese preference, but a "ceiling effect" created by infants' overall very high level of interest in the videos (they looked at the videos 93% of each trial on average) may have undercut the sensitivity of our method to such a preference. In support of this more nuanced perspective, we found that infants displayed longer average looking to motionese relative to adult-directed videos for five of the six videos, and we also found a trend for motionese looking to increasingly exceed adult-directed action-looking across blocks as the study proceeded, presumably because the ceiling effect progressively attenuated. In Chapter IV we discuss further possible causes of a ceiling effect and thoughts about how to ameliorate it in future studies.

It is of course also possible that our methodology was insensitive to a motionese preference in other respects. For example, it is worth noting that the method used in the original study (Brand & Shallcross, 2008) demonstrating that infants prefer to look at motionese over adult-directed action was a preferential looking task. Infants were shown two videos, one adult-directed and the other infant-directed, and experimenters assessed which of the two videos infants looked to more. In contrast, we showed infants only one

video at a time. Presenting a single stimulus was necessitated by the range of research questions we were addressing in this research. However, it is nevertheless possible that the preferential looking method is better suited to detecting such a preference than the single stimulus method we employed. Likewise, had we used an infant-controlled version of the single-screen procedure (used in some research documenting infants' preference for motherese over adult-directed speech; e.g., ManyBabies Consortium, under revision) rather than the fixed-timing version we opted for, we also might have observed a more systematic motionese preference.

Another collection of analyses addressed our second and third questions: whether infants would display a PDR in response to major action boundaries, and whether motionese would facilitate such boundary-related responding. As it turned out, a significant boundary effect – increased pupil diameter during boundary regions – was observed, but only for motionese demonstrations. These findings suggest that motionese indeed enhance infants' detection of segmental structure in unfolding activity. It is worth noting, however, that infants' boundary-related PDR differed in other ways from the comparable pattern observed in one previous study with adults. For one thing, infants' boundary-related PDR during motionese demonstrations was slower (occurring on average within a one-second region after the boundary) than adults' (occurring on average within a half-second region after the boundary). This was consistent with other evidence that infants' pupil response is generally slower than adults' (e.g., Verschoor et al., 2013, 2015; Zhang et al., 2019). Another difference was that a linear trend provided the best characterization of infants' boundary-related PDR; in contrast, while Tanaka and colleagues (in preparation) did find both significant linear and quadratic trends, the

quadratic trend was stronger and suggested that adults' pupil diameter began to return to baseline shortly after their boundary-related PDR. Why might infants' pupil diameter remain high after the boundary? One likely explanation is again, that infants' pupil response – including the return to baseline – may simply be slower than adults'. Also, as we observed from visual examination of the videos, there was often considerable post-boundary movement in infant-directed demonstrations. Upon examination of the videos, we observed that caregivers depicted in the videos frequently did things like spreading their arms to exaggerate the fact that a boundary had occurred. Often caregivers would also make excited and exaggerated facial expressions after finishing a unit of action. These features of the stimuli are of course characteristic of motionese, and could serve to sustain infants' arousal, thereby reducing a tendency for pupil diameter to return to baseline levels after a boundary. In an exploratory analysis, we found some evidence that the boundary effect emerges across time, with the boundary-related PDR gaining systematicity after several viewings. This is consistent with previous research conducted in our lab (e.g., Kosie & Baldwin, 2019a) and suggests possible reorganization of attention across repeated viewing of novel activity sequences.

Regarding luminance, we considered the possibility that luminance patterns might have influenced our results. While we took a number of steps to control for luminance across our videos (described in detail in Chapter II), luminance for motionese demonstrations was high pre- and post-boundary, and lower within the boundary region. In the pupillometry results, we observed larger pupil diameter during the boundary region and smaller pupil diameter during the pre-boundary region. These PDR results would generally correspond to the observed patterns of luminance; however, there were a few

reasons why luminance did not seem to offer a sole explanation for differences in pupil size across regions. For one, the change in luminance across our videos was smaller than would be expected to significantly influence an observers' pupil diameter (Bala, unpublished data). Additionally, we directly controlled for luminance in many of our analyses. Luminance itself was not a significant predictor of pupil size, nor did controlling for luminance influence the relationship between video region (pre-boundary, boundary, post-boundary) and pupil size. Finally, as mentioned previously, luminance was larger pre- and post-boundary than at the boundary itself, but the biggest difference in luminance was the difference from boundary to post-boundary. In contrast, we found that infants' pupil size increased at the boundary and remained high post-boundary. If luminance alone influenced infants' pupil diameter, we would expect a constriction post boundary, and this was not observed. Thus, there is strong evidence that our pupil-related effects cannot be explained simply by video luminance.

Taken together, the results of these analyses suggest that motionese promotes infants' detection of major action boundaries within unfolding action. While we did not find an overall motionese preference reflected in infant looking or infant pupil diameter, we believe that the current stimuli were not well-suited to find this effect. Suggestions for changes to be implemented in future research are discussed in Chapter IV. However, while infants' PDR was not indicative of enhanced attention to action boundaries in adult-directed activity sequences, a boundary advantage (increased PDR at action boundaries) emerged when they viewed motionese action. In addition to providing the first demonstration that motionese enhances infants' online action processing, this

research further validated the use of a new, inexpensive, open-source, and infant-friendly methodology for measuring infants' attention to streaming visual stimuli.

CHAPTER IV

GENERAL DISCUSSION

In this dissertation, we used a new pupillometry technology (the SIPR PDR system) to address several questions about infants' response to motionese, the modified form of action that adults engage in when demonstrating novel object properties to infants. Unexpectedly, the pupillometry findings revealed only weak evidence that infants prefer motionese over adult-directed action; however, infants' overall high level of interest in all action demonstrations likely undercut detection of a previously documented motionese preference. Of particular interest, infants displayed a prominent pupillary response to action boundaries within continuously unfolding activity, but only in the context of motionese demonstrations. This finding provides the first evidence to date that motionese action modifications alter infants' online action processing. In particular, motionese scaffolds infants' detection of segmental structure within dynamically unfolding action.

A corpus of infant- and adult-directed action

A first step in this work was to locate a set of videos that matched a number of specific criteria we set for our research. We hoped to acquire a set of videos that contained short clips of caregivers demonstrating novel object properties to both their infant and an adult interaction partner. Desired design features included (1) each clip containing one major action boundary, with the temporal location of the boundary varying across demonstrations of different objects, (2) naturalistic action that

nevertheless maximized the difference between infant- versus adult-directed demonstrations, and (3) balanced luminance across adult and infant-directed action, as pupil size is influenced by luminance (e.g., Loewenfeld, 1993). Because no existing set of videos met all (or even most) of these criteria, we opted to collect a new video corpus.

For this work, and future work with these videos, we collected additional information about infants' knowledge and response to the objects employed, as well as about the caregivers featured in the videos and their infants. To validate the objects used, we assessed infants' interest in the objects in multiple ways as well as gathering information from caregivers regarding infants' familiarity with each of the objects demonstrated. Additional available information about each infant and family included basic demographic information (e.g., gender, race, socioeconomic status, etc) and a measure of infants' early language ability. Though we presented the videos silently, the original digital video files contain both audio and visual information, enabling future investigation of caregivers' speech as well as action to adult and infant partners. Finally, for some participants, we also had a video camera directed at the infant during demonstrations. In future work, these videos can be synchronized with adult demonstrations and infants' behavior can be coded. Thus, in addition to providing stimuli for use in the current research, we have produced a large corpus of adult -and infant-directed action that provides a resource for potential future work addressing a host of questions about the nature of caregiver/infant interactions.

Comparison of infants' interest in motionese versus adult-directed action

To explore infants' overall processing of motionese versus adult-directed action, we collected data from another sample of infants as they viewed a carefully selected subset of the video stimuli just described. In particular, we analyzed both infants' looking and their pupil size in response to the motionese and adult-directed demonstrations depicted in this subset of videos. In contrast to our predictions, we did not find a significant difference in looking or pupil size to motionese versus adult-directed demonstrations. However, we did observe a number of hints of a motionese preference, mitigating what might otherwise seem to be a non-replication of previous findings. In particular, (1) infants exhibited longer looking on average to motionese than adult-directed demonstrations on five out of six blocks of trials, (2) the trend to look at motionese over adult-directed action increased across blocks, and (3) infants' average pupil size tended to be larger in response to motionese over adult-directed activity sequences, though none of these effects quite reached statistical significance. The failure to reach significance appears to have been strongly influenced by a "ceiling effect" in infants' looking. On average, infants were looking for 93% of each trial (93% for adult-directed and 94% for infant-directed trials), perhaps inhibiting detection of a difference in looking across the two trial types. This high level of interest in both trial types, regardless of the identity of the interaction partner, may additionally have impacted infants' pupil size (which reflects a general level of interest and arousal), potentially obscuring detection of possible differential interest in motionese versus adult-directed action with this measure as well.

The primary consideration driving our choice of stimuli for the pupillometry study involved ensuring that each video contained one major action boundary and that the

boundary was aligned as best as possible across video pairs (i.e., infant- and adult-directed versions of the same actor interacting with the same object). Perhaps our focus on balancing videos with respect to these characteristics minimized differences between motionese and adult-directed videos, potentially reducing the possibility of looking-time differences emerging for these two types of videos. As well the videos were short, ranging from seven to twelve seconds in length. Our rationale for the short clips was to (1) enable us to find segments featuring the same action on the same object across infant- and adult-directed demonstrations, (2) remove extraneous activity (that differed across motionese and adult-directed demonstrations) occurring before and after these matched segments of action, (3) capture only one major action boundary, and (4) be short enough that infants could complete multiple viewings of the same clip, to best estimate and aggregate infants' pupillary response to that particular activity sequence. Together, these design features likely contributed to infants' high level of interest in all videos and could be modified in future research.

Additionally, in our effort to equate the videos, we may have removed some important features that typically distinguish motionese from adult-directed action. This is reflected in results from the motionese coding described in Chapter II. While infant-directed videos were rated significantly higher than adult-directed videos on interactiveness and enthusiasm, they differed less on dimensions such as range of motion, rate, and repetitiveness, which have been found to be enhanced in motionese demonstrations in previous research (e.g., Brand, et al., 2002). On the one hand, the pattern of ratings we observed were not unexpected. For example, we specifically designed the pairs of infant- and adult-directed videos to feature the same action on the

same object, be equated in duration, and feature one major action boundary that occurred at approximately the same moment across videos. We thus knowingly eliminated any possible difference on dimensions such as repetition and rate. On the other hand, it is possible that, if videos had preserved all of the features that distinguish motionese from adult-directed action in prior research, the predicted differences in looking time and pupil size between infant- and adult-directed demonstrations would have emerged as statistically significant.

Finally, we measured infants' looking to infant- and adult-directed videos presented one at a time for a fixed duration. In contrast, in the original work demonstrating that infants prefer to view motionese over adult-directed demonstrations, researchers used a preferential looking paradigm (Brand & Shallcross, 2008). That is, infants were presented with an infant-directed video on one side of the screen and, at the same time, an adult-directed video on the other side of the screen. The amount of time infants spent looking to each side of the screen was coded as a measure of their preference for motionese versus adult-directed action. Using our stimuli, this might have been a better way to address whether infants exhibited a preference for viewing motionese versus adult-directed action, even though their baseline interest in both was high. Against this, other studies (e.g., ManyBabies Consortium, under revision) investigating infants' preference for motherese speech relative to adult-directed speech have presented infants with infant- and adult-directed speech in a fashion similar to the method we used (i.e., infants heard *either* infant- or adult-directed speech, and time spent looking to a central fixation stimulus was measured as an index of infants' preference), yet they found that infants preferred to listen to infant-directed speech over adult-directed

speech. One critical difference between their method and ours, however, was that their procedure was infant-controlled for the majority of participating labs. Rather than exposing infants to infant- and adult-directed speech for a fixed duration of time, exposure stopped when infants looked away for more than two seconds. We intentionally avoided the infant-controlled design in our procedure, primarily because an important goal of our design was to ensure that infants had the opportunity to see the major action boundary occurring within each video in order to test whether they displayed a PDR to the boundary; on an infant-controlled procedure they might have been more likely to miss viewing the boundary. As it turns out, however, evidence amassed in the large-scale, multi-site ManyBabies Consortium study (under revision) suggests that infant-controlled procedures, and especially procedures requiring more effort from the infant (e.g., head-turn preference procedures), may be more sensitive to a motherese preference than non-infant-controlled procedures. Thus, the fact that infants in our study had no control over their exposure time to infant- and adult-directed action may have additionally undercut the sensitivity of our measure and our ability to detect a strong systematic preference for motionese over adult-directed action as has previously been observed.

Motionese facilitated infants' ability to find structure in unfolding action

We found that infants' pupil size increased in response to boundaries in motionese, but not adult-directed, action. These results supported our prediction that motionese highlights structure within activity as it unfolds across time. This finding raises an obvious next question: Precisely what is it about motionese that facilitates infants' detection of action boundaries? One hypothesis is that it's something about the

demonstration itself – perhaps caregivers move and manipulate objects in ways that highlight structure within dynamic activity. Another possibility is that motionese simply heightens infants' attention, which increases the chances that they will detect structure within unfolding activity. Perhaps instead, or in addition, motionese indicates to the infant that this demonstration is "for me." When infants can infer that an action demonstration is directed to them, this might further enhance their attention and thus facilitate their detection of structure as activity unfolds. Evidence from the current study speaks to all of these interpretations.

First, could something about the motionese demonstrations have enhanced infants' detection of action boundaries? In the related domain of motherese, or infant-directed speech, it has been suggested that specific characteristics of motherese input promote infants' ability to find structure in speech (e.g., Kemler-Nelson et al., 1989; Gleitman, Newport, & Gleitman, 1984). Perhaps characteristics of motionese similarly facilitate infants' detection of structure within dynamic activity. As described previously, however, the steps we took to match infant- and adult-directed demonstrations reduced some characteristics of motionese that might otherwise serve to highlight action boundaries. For example, shorter action sequences – often characteristic of motionese – might highlight boundaries with pauses or repetition of shorter units of action. However, these dimensions of motionese were reduced in our stimuli. We did find, though, that pixel values – sometimes used as an index of motion change (e.g., Hard et al., 2011; Loucks & Baldwin, 2009) – were greater both before and, after, action boundaries in motionese demonstrations. Visual examination of our videos confirmed that this large degree of pixel change often corresponded to body movements that might highlight the

fact that a boundary had just occurred (such as large, emphatic arm movements). Additionally, enthusiasm and interactiveness were high in our infant-directed demonstrations. There is some evidence (e.g., Brand et al., 2013) that caregivers' gaze toward infants, reflected in the "interactiveness" coding, coincides systematically with action boundaries. It is additionally possible that exaggerated facial expressions, which likely contributed to higher observed enthusiasm ratings in our findings, frequently coincided with action boundaries. These features of our motionese demonstrations could have facilitated infants' detection of structure. Further coding of the video stimuli will be necessary to fully explore these possibilities; this represents an interesting future follow-up to the dissertation research.

A second hypothesis is that motionese increases infants' attention overall and, if infants' attention is increased, they might be better able to attend to action and thus to detect segmental structure. Prior research supports this hypothesis: when infants are in an attentive state (as indexed by heart rate) during stimulus presentation, they are more readily able to recognize that stimulus at later test (Richards, 1997; Frick & Richards, 2001). While that research focused on infants' recognition memory, and not their sensitivity to structure as in the present dissertation, there is reason to believe that the two might be related. For example, a substantial body of evidence suggests that infants' (and adults') attention to structure within action is linked to later memory (e.g., Sonne et al., 2016, 2017; Hard et al., 2011; Zacks et al., 2006). While we did not find a significant difference in infants' overall attention to infant- over adult-directed action, there were a number of hints that a motionese preference was at least weakly present. Thus, despite the fact that these comparisons did not reach statistical significance, infants may have

been in a more attentive state in response to motionese demonstrations, enhancing their processing of the unfolding activity.

The final alternative we've suggested above is that motionese indicates to infants that this demonstration is "for me." Information presented to infants in a social context appears to facilitate learning (e.g., Baldwin, Markman, Bill, Desjardins, Irwin, & Tidball, 1996; Baldwin, 2000; Akhtar & Tomasello, 2000; Sage & Baldwin, 2011; Csibra & Gergely, 2009), which seems to be either illustrative of, or closely related to, a phenomenon that Kuhl and colleagues (e.g., Kuhl, Tsao, & Liu, 2003; Kuhl, 2007) call "social gating" following a similar phenomenon in bird-song learning (e.g., Doupe & Kuhl, 1999; Kuhl, 2003). One interpretation of social gating is that a social context simply elicits an increase in infants' overall attention, analogous to our second alternative account outlined above. However, it has been demonstrated that infants presented with stimuli in both social and non-social contexts learn better from the social context, despite equivalent attention to stimuli across contexts (e.g., Baldwin et al., 1996; Sage & Baldwin, 2011). Thus, there is likely to be something more driving infants' learning from social stimuli like the motionese demonstrations in the current research. Perhaps contributing to this effect, Gergely, Csibra, and colleagues (Csibra & Gergely, 2006, 2009, 2011; Gergeley, Egyed, & Kiraly, 2007) suggest that pedagogical cues, which abound in motionese, signal to infants that they are being taught and, consequently, infants adopt a "pedagogical stance" that primes them to learn. Along these lines, in our current work one of the features displaying the greatest differential across infant- versus adult-directed demonstrations was interactiveness; infant-directed demonstrations were rated much higher on this dimension than adult-directed demonstrations. Interactiveness

involves gaze toward the interaction partner and bids for joint attention, cues that are proposed to be key signals of natural pedagogy. Perhaps, then, motionese promoted infants' adoption of a pedagogical stance, and thereby enhanced their detection of segmental structure in unfolding activity sequences.

The current study provided evidence consistent with all of these alternative accounts, without singling out any particular account as the most plausible mechanism by which motionese could enhance infants' attention to structure. At this juncture, it seems unlikely that any one of the mechanisms proposed above can fully explain why infants displayed a pupillary response to action boundaries within infant-directed demonstrations but not to comparable boundaries within adult-directed action. In contrast, it seems plausible, and perhaps even likely, that all these mechanisms operated in concert to enhance infants' processing of dynamic action.

Limitations

While this research provides altogether new information about the influence of motionese on infants' processing of everyday activity, we note several limitations. The first concerns our ability to interpret infants' increased pupil diameter as a response to action boundaries, *per se*. Supporting this interpretation, we found a systematic PDR in motionese activity sequences that occurred within a one-second window after action boundaries, and this effect was stable even when controlling for frame-by-frame pixel values (which reflect luminance of the stimulus as well as, at least to some degree in our stimuli, motion change occurring in the activity sequence). However, this conclusion relies heavily on the fact that the one-second window is an appropriate region in which to

expect infants' boundary-related PDR to occur. While this was based on (limited) prior evidence and our own investigations into the timing of infants' pupillary response, further validation of this response window is necessary to increase confidence in our findings. Additionally, further research is needed to confirm that infants are indeed responding to a boundary and not extraneous features of the activity sequence. Suggestions for ways to investigate these questions are outlined in the Future Directions section below.

We made efforts to ensure that the stimulus videos in the current study were both naturalistic and representative representations of infant- and adult-directed action, but it is unlikely that our videos fully captured the nature of infants' everyday experience. First, we intentionally selected videos from the corpus of caregiver-infant interactions that featured very distinct and obvious differences in infant- versus adult-directed action. Thus, the mothers featured in these videos might engage in greater than average levels of motionese. It is unlikely that all infants receive such distinctly different formats – if an infants' caregiver doesn't use a lot of motionese, the differential between infant- and adult-directed action in their everyday input might not be as pronounced. It is unclear what effect the large differential between infant- and adult-directed action might have had on infants' processing. Another focus in stimulus creation was to align boundaries across infant- and adult-directed demonstrations and to equate these demonstrations to the extent possible. Consequently, features of motionese such as repetition and simplification were likely less prevalent in our stimuli than in the real-world action to which infants are exposed. Additionally, to enable the collection of pupillometry data, our stimuli were videos rather than live demonstrations. Thus, certain other features of motionese – such as opportunity for frequent object exchanges – were not available to infants. To isolate

the influence of infant- versus adult-directed *action*, and not speech, our videos were presented in silence. This, too, is likely a stark contrast to infants' everyday experience in which interactions with adults often consist of coordinated action, speech, touch, and other social sources of information. In sum, while our coding of the videos increased confidence that the infant-directed demonstrations did indeed contain features characteristic of motionese, there are limitations to broad generalizability of our results. That said, many of these limitations would seem, on the whole, to have been likely to have reduced the chances that we would detect benefits of motionese on infants' on-line action processing. Nonetheless, we indeed observed such benefits.

In addition to features of our stimuli, some characteristics of our participants themselves engender limitations to generalization. For example, the participants in both the corpus creation project and the pupillometry study were highly educated with little variability in socioeconomic status (SES). The extent to which motionese is present in the input of lower SES infants is currently unknown, nor is it known how infants from a lower SES demographic might respond to action containing features of motionese. However, there is evidence that socio-economically disadvantaged children are at risk for cognitive and linguistic deficits (e.g., Neville, Stevens, Pakulak, & Bell, 2013) and may receive lower quality input more generally (e.g., Hoff, 2003; Bettes, 1988). Thus, an important next step would be to replicate this research with a lower SES sample and to explore the efficacy of motionese as an intervention for children who are at-risk. Our sample was similarly homogeneous with respect to race/ethnicity – *all* participants in the pupillometry study and nearly all participants included in the corpus identified as white (for at least one of the races they selected). While there is little evidence regarding the

extent to which motionese is present across racial/ethnic groups and cultures (but see Gogate, Maganti, & Bahrick, 2015; Kline, Boyd, & Henrich, 2013; Kline, 2015), there is evidence in the language domain that most, but not all, cultures use motherese speech (e.g., Blount & Padgug, 1976; Ferguson, 1964; Schieffelin, 1979; Fernald , Taeschner, Dunn, & Papousek, 1989). While it seems likely that motionese would similarly be found across a variety of cultures, it is unknown whether the results of this research would replicate outside of a North American, English speaking, white, higher SES sample.

Though we had a moderately large infant sample for corpus creation (N = 53), the difficulties of recruiting a developmental population in a more restricted age range (9-12 months) resulted in a relatively smaller sample (N = 27) for the pupillometry study. Thus, we plan to continue collecting data to attain a larger sample size before submitting this work for publication. Despite this, a sample size of 27 infants is within the range that is typical across infant pupillometry research (e.g., Sirois & Jackson, 2011; Jackson & Sirois, 2009; Verschoor et al., 2015), and is larger than the minimum sample size that Oakes (2017) suggests for infancy research more generally (she suggests, at minimum, N = 24, though this was estimated using simulations from published looking-time studies rather than pupillometry data). Additionally, although the pupillometry sample is relatively small, we succeeded in collecting considerable data from each individual infant, with pupil size measured throughout a median of 29 trials per infant (and a total of 696 trials across the sample of all infants). The large amount of data obtained from each infant provided a strong estimate of within-subject effects and thus increased our available power (e.g., DeBolt, Rhemtulla, & Oakes, 2019). Still, data from a larger sample will further increase statistical power, enabling a more robust estimate of the

extent to which motionese influences infants' processing of dynamic activity. Future high-powered replication of these results will of course also be valuable in providing further information regarding what appears from the present research to be a facilitative role of motionese on infants' action processing.

Broader Implications

On its own, the video corpus that we created to generate the stimuli for this research provides a substantial contribution to the study of caregiver-infant interaction. The creation of this corpus facilitates investigation of a host of new research questions regarding the nature of infant-directed language and action (described in further detail below in the Future Directions). Also, because all videos will be archived on Databrary (with caregivers' consent; Databrary, 2012), this corpus provides an opportunity for a diverse group of researchers to address a variety of questions about the nature of caregiver-infant interaction. In fact, the corpus has already garnered interest from robotics researchers seeking to design computational systems that incorporate features of motionese into their child-directed action demonstrations. We are hopeful that other researchers will also both use, and contribute to, this corpus of infant- and adult-directed action, increasing both its size and diversity.

This dissertation research also validates a novel, open-source, inexpensive, infant-friendly pupillometry technology, offering another important contribution to developmental science. In particular, the dissertation findings provide the first demonstration that the SIPR technology can be used for measuring infants' processing of streaming visual information. This opens up a new landscape of potential research. Both

inexpensive and portable, the SIPR system is potentially accessible to diverse research labs, amenable for working with challenging populations such as infants, and it can used in locations such as preschools and children's museums, thereby enabling the collection of very large samples and substantially increasing statistical power for future pupillometry studies. Thus, validation of this methodology opens up opportunities to address any number of questions about infants' attention to streaming visual stimuli and to increase the power and diversity of developmental pupillometry research.

Perhaps the most exciting advance resulting from validation of this methodology is the window it can provide on learning as it unfolds over time. With continuing refinement, our hope is that we can use this pupillometry system can be employed across many research contexts to observe changes in infants' attentional patterns as they first encounter novel stimuli and learn the structure of their input across repeated exposure. The ability to watch as learning unfolds would provide insight into diverse influences on infants' processing of novel input. Further, having a window on infants' processing as it occurs in real time will enable exploration of individual differences in infants' attentional allocation to stimuli and what these differences predict about infants' learning. As a result, we may be able to identify infants who are at risk for learning challenges and to develop systems for early intervention.

Finally, this dissertation extends current understanding of the ways in which motionese benefits infants' development. Previous research has documented the motionese phenomenon, that human caregivers spontaneously modify motion when demonstrating action to infants. As well, prior work demonstrated both that (1) infants prefer motionese over adult-directed action, and (2) motionese promotes infants'

imitation of novel activity sequences. What had remained mysterious, however, was the precise ways in which motionese might alter infants' processing of dynamically unfolding activity. This question had been difficult to address, in part, because existing methodologies were not well-suited to probing infants' moment-to-moment action processing. The research reported in this dissertation offers a signal advance on this methodological front, and at the same time provides the first evidence to date that motionese promotes infants' detection of segmental structure within dynamically unfolding activity. Put another way, the current findings indicate that, by providing motionese demonstrations, caregivers spontaneously enhance infants' detection of boundaries within continuous activity. This facilitates infants' discovery of action units that are amenable to encoding in memory for later recall, and likely promotes their efficient processing of similar activity sequences when subsequently encountered.

Future directions

In this research we addressed three questions about infants' processing of activity as it unfolds across time, focusing on (1) the influence of motionese on infants' overall attention to action, (2) infants' response to boundaries within continuously unfolding activity, and (3) motionese as a mechanism for scaffolding infants' processing of dynamic action. While this work provides the first insight into infants' online processing of activity and caregivers' influence on this processing, some issues remain as yet unaddressed on each of these points. Additionally, in carrying out this research a number of methodological questions arose that point to the need for future investigation into best practices for working with infant pupillometry data.

Methodological questions raised in this research. A basic methodological question concerns the timing of infants' PDR to action boundaries. In the current work, to determine the appropriate window in which to examine infants' pupil size for boundary-related effects, we consulted previous research and examined the timing of infants' response to a perceptual/cognitive event (i.e., the initiation of movement in a video) in our own data. However, a more systematic investigation into the timing of infants' PDR to cognitive stimuli would increase confidence regarding the appropriate window within which to explore infants' response to action boundaries. Additionally, as yet, no direct, systematic comparison of the timing of infants' and adults' PDR has been undertaken, and this too marks an important future direction for this work. An experiment comparing PDR across infants and adults could be as simple as exposing observers to a cognitive event such as the appearance of a stimulus (or a variety of different such cognitive events), and comparing the time-course and magnitude of infants' versus adults' PDRs. The results of such comparisons would provide useful information for designing a variety of future research studies.

More broadly, when reviewing the existing infant pupillometry literature, we found striking diversity in procedures for data collection, preprocessing, and analysis. This highlighted the need for methodological investigation into the consequence of these diverse practices for working with pupillometry data. A valuable next step could be an investigation of the extent to which the use of different methods across research laboratories influences the findings of pupillometry research. In recent work with adults, for example, researchers provided teams of analysts with a single dataset and asked them to separately test the same research question (Silberzahn et al., 2018). They found marked

variability in results, highlighting the influence of lab-specific analytical decisions on research outcomes. A similar study might be informative here – a set of labs who use pupillometry with infants could be provided with a single dataset reflecting the results of a study with a very simple research design (such as infants' PDR to a stimulus appearing on a screen). Researchers from each lab could analyze the data using their typical approach, and the results could then be compared across labs. The results of this work would provide important insight into the consequences of cross-lab heterogeneity in pupillometry methods. Our hope is that the results of this collaborative research would additionally inform best practices for working with at least certain types of infant pupillometry data and prompt researchers to take seriously the consequence of decisions related to methods of data collection, preprocessing, and analysis.

The influence of motionese on infants' overall attention to unfolding activity. In the current research, we found weak evidence that infants' attention was enhanced by motionese (relative to adult-directed action). We suggest that this failure to find a significant difference in infants' looking to motionese over adult-directed action seems to be due to a ceiling effect. Above, we describe how certain features of our stimuli might have driven this overall high level of attention to both motionese and adult-directed action. That said, it is important to recognize that no systematic exploration has yet been undertaken regarding the factors that influence infants' response to motionese. An open question concerns how diversity across infants' everyday input relates to their processing of motionese activity. As mentioned in our Limitations section, some groups of children – such as those in lower SES families or with caregivers who are suffering from depression – might encounter motionese less frequently or in attenuated form. It is an

open question how such differences influence the development of infants' fluency in action processing. Motionese input might be particularly salient if infants have not seen such modified action in the past, and therefore particularly influential in supporting their detection of structure within unfolding activity. On the other hand, if infants haven't regularly experienced motionese, it might seem foreign and perhaps overly stimulating, thus undercutting efficient processing of the unfolding action. In the future, it would be informative to describe the quantity and quality of motionese input infants receive from their caregivers and examine the extent to which characteristics of this input relates to infants' preference for, and pupillary response to, infant- versus adult-directed action. The results of this work could have important consequences for fully understanding the potential of motionese to help at-risk infants learn to process everyday activity.

The video corpus created as part of this dissertation work opens up a number of avenues for possible future research describing the nature of motionese input more generally. For one, this corpus holds potential to promote understanding of when and why caregivers use motionese. Questions that can be addressed with the corpus include: Are caregivers more likely to use motionese when they think an object is more novel to their infant? What aspects of infants' behavior (such as interest or responsiveness) are correlated with caregivers' tendency to increase or decrease their use of motionese? Additionally, we can use the corpus to investigate the extent to which motionese relates to other dimensions of natural pedagogy. One set of questions we intend to explore centers on the extent to which motionese and motherese are correlated and/or complementary phenomena: Are caregivers who use more motionese also likely to use more motherese? Are there times at which caregivers rely on one versus the other?

Investigations of these questions will enhance overall understanding of (1) how caregivers modify their behaviors in ways that help infants learn, as well as (2) the extent to which caregivers differ in their use of natural pedagogy.

In addition to motionese, a variety of other factors hold potential to influence and facilitate infants' processing of dynamic visual stimuli. For example, there is evidence that infants' own action experience influences later action perception (Sommerville et al., 2005). However, precisely how such action experience alters infants' processing of unfolding activity is not yet known. It seems plausible that enhanced detection of action boundaries within the motion stream may be one aspect of processing that is benefitted by action experience. This prediction has not as yet been explored, but is amenable to test with the pupillometry method validated in this research. Additionally, factors including complexity of input (e.g., Dawson & Gerken, 2009), variability (e.g., Gomez, 2002), and context (e.g., Roy, Frank, DeCamp, Miller, & Roy, 2015), systematically relate to infants' learning from their environment. However, how these factors influence learning has often been tested by first exposing infants to stimuli and attempting to infer what they learned at later test. The pupillometry method used in this work holds promise for elucidating just how these factors influence infants' online processing and how that relates to what they learn from various sources of input.

Exploring infants' processing of action boundaries in the absence of motionese. In the absence of motionese, infants in this dissertation research did not exhibit a boundary-related PDR. However, a substantial body of research suggests that, even without motionese, infants are indeed sensitive to action boundaries in at least some kinds of activity (e.g., Baldwin et al., 2001; Hespos et al., 2009, 2010; Saylor et al., 2007;

Roseberry et al., 2011; Stahl et al., 2014; Monroy et al., 2017). Why, then, did we not observe a PDR to action boundaries for adult-directed action? Are there types of adult-directed activity in which we *would* observe a boundary-related PDR? This is an important direction for future research. One hypothesis is that infants failed to show a boundary-related PDR while observing adult-directed activity simply because it was presented in the context of alternation with infant-directed action. Perhaps something about the infant-directed demonstrations was more salient and drew infants' attention away from the adult-directed versions, thus altering their processing. Alternatively, we hypothesized earlier that the pedagogical context led infants to infer that the infant-directed versions were "for me." Correspondingly, they may have inferred that the adult-directed versions were not "for me," and thus paid less attention to these activity sequences. An obvious future direction, then, is to present infants only the adult-directed versions and explore the patterns of their PDR in the absence of an infant-directed comparator. On the other hand, the infant- and adult-directed sequences we presented to infants were quite novel, as verified in Chapter II (i.e., caregivers' ratings of how likely it was that infants had come in to the session knowing what to do with the objects were low across the board). The actions involved in prior research investigating infants' processing of dynamic activity were relatively more familiar, such as an actor picking up a towel from the kitchen floor. If we had presented action that was more familiar to infants, we might have also found boundary-related PDR during viewing of adult-directed activity. This difference between novel and familiar activity would suggest that motionese might be particularly helpful when infants are encountering actions for the first time and less important when action is already familiar.

Further investigation of infants' boundary-related PDR to motionese activity sequences. In infant-directed activity sequences (e.g., motionese), infants exhibited a boundary-related PDR. However, as mentioned previously, an open question is whether this PDR was indeed in response to action boundaries. For example, features unrelated to the boundary could influence infants' pupillary response, and it is possible that these features just happened to occur more often at boundary regions in motionese videos. Systematic coding of our video sequences might serve to isolate infants' boundary-related responses from responses to extraneous features occurring in the temporal region that coincides with action boundaries. For example, in future research it would be useful to develop a coding scheme for activity that regularly occurs in synchrony with action boundaries and determine the frequency with which these activities occur at boundary and non-boundary regions. This would enable us to (1) quantify the extent to which these activities occur only at boundary regions, (2) investigate how well these activities alone (regardless of whether they occur at boundaries) are predictive of infants' PDR, and/or (3) control for these activities in pupillometry analyses.

A statistical-learning approach would enable us to address whether one and the same juncture in an activity sequence could be identified as both boundary and non-boundary regions, depending on infants' knowledge state. In a statistical-learning paradigm (e.g., Baldwin et al., 2008; Hard et al., 2018; Roseberry et al., 2011; Stahl et al., 2014) observers learn the structure of a stimulus over time. For example, as in prior research, infants might be presented with actions that have underlying statistical regularities (e.g., *poke* always follows *pour* while *drink* follows *poke* only a third of the time). Given time to learn these learn these regularities, we would expect actions that

regularly co-occur to cohere into larger action units with boundaries between them (i.e., *pour-poke* is a unit, but there would be a boundary between *poke* and *drink*). With repeated exposure to the sequence over time, we might expect to see a boundary-related PDR emerge. Specifically, when infants are first exposed to the stimuli, the juncture between *pour* and *poke* should be processed similarly to the juncture between *poke* and *drink*. However, once they've learned that *pour-poke* is a statistically-coherent unit, a systematic boundary-related PDR should emerge only for the transition between *poke* and *drink* and not for the transition between *pour* and *poke*. This finding would help to clarify that the PDR patterns we are attributing to infants' identification of action boundaries within motionese indeed reflect the identification of action boundaries, *per se*. Additionally, a statistical-learning approach would rule out the issue of coinciding features being the sole explanation of infants' increased PDR at boundary regions; one and the same temporal region within an activity sequence would serve as both a boundary and non-boundary as the statistical structure is learned.

To investigate the extent to which our observed PDR occurs systematically at action boundaries, a final direction for future research would be to engage in a series of simulations. While we have not yet outlined the precise details involved in carrying out such simulations, we are currently discussing the possibility of such an investigation in relation to our pupillometry work with adults (i.e., Tanaka et al., in preparation). In this series of simulations, we might repeatedly shuffle our pupil size values and explore the frequency with which these shuffled values would show a boundary-related PDR simply by chance. If the likelihood that a boundary-related PDR would emerge simply by shuffling the pupil size values is low, it would support the interpretation that infants' (and

adults') pupils dilate specifically in response to the occurrence of action boundaries or at least to activity occurring specifically at boundary regions.

Conclusion

In conclusion, this dissertation makes several important contributions to developmental science. To conduct this research, we created a large video corpus of infant- and adult-directed action and language. In the future, we intend to use this corpus to further understanding of motionese in particular, and natural pedagogy more broadly. This set of videos will be made open to other researchers who might be interested in using our corpus to advance knowledge of the dynamics of caregiver-infant interaction.

This dissertation also validated a new, open-source pupillometry technology for investigating infants' processing of streaming visual stimuli. With this methodology, we demonstrated that infants displayed a systematic increase in pupil size in response to action boundaries within sequences of novel activity, but only when that activity was in a motionese format. This finding offers altogether new insight into precisely how motionese benefits infants' ability to find structure in action as it is unfolding in time.

REFERENCES CITED

Addyman, C., Rocha, S., & Mareschal, D. (2014). Mapping the origins of time: Scalar errors in infant time estimation. *Developmental psychology*, *50*(8), 2030.

Akhtar, N., & Tomasello, M. (2000). The social nature of words and word Learning. In R. Golinkoff & K. Hirsh-Pasek (Eds.), *Becoming a word learner: A debate on lexical acquisition*, Oxford, U.K.: Oxford University Press.

Ambrosini, E., Reddy, V., De Looper, A., Costantini, M., Lopez, B., & Sinigaglia, C. (2013). Looking ahead: anticipatory gaze and motor ability in infancy. *PloS one*, *8*(7), e67916.

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological science*, *9*(4), 321-324.

Aston-Jones, G., & Cohen, J. D. (2005). Adaptive gain and the role of the locus coeruleus–norepinephrine system in optimal performance. *Journal of Comparative Neurology*, *493*(1), 99-110.

Bailey, H. R., Kurby, C. A., Giovannetti, T., & Zacks, J. M. (2013). Action perception predicts action performance. *Neuropsychologia*, *51*(11), 2294-2304.

Bakeman, R., & Adamson, L. B. (1984). Coordinating attention to people and objects in mother-infant and peer-infant interaction. *Child development*, 1278-1289.

Bala, A. (in preparation). [Pupil diameter in response to screen luminance.] Unpublished data.

Bala, A., Keller, C., Whitchurch, E., Baldwin, D., & Takahashi, T. (2016, October). Pupillary dilation as a hearing screening in adults and infants. Poster presented at the 2016 Northwest Auditory and Vestibular Research Meeting, Portland, OR.

Baldwin, D.A. (2000). Interpersonal understanding fuels knowledge acquisition. *Current directions in psychological science*, *9*(2), 40-45.

Baldwin, D.A. (2012). Redescribing action. In Banaji, M. R. & Gelman, S. A. *Navigating the social world: What infants, children, and other species can teach us.* New York: Oxford University Press.

Baldwin, D.A., Andersson, A., Saffran, J., & Meyer, M. (2008). Segmenting dynamic human action via statistical structure. *Cognition*, *106*(3), 1382-1407.

Baldwin, D. A., Baird, J. A., Saylor, M. M., & Clark, M. A. (2001). Infants parse dynamic action. *Child development*, *72*(3), 708-717.

Baldwin, D. A., & Kosie, J. E. Intersubjectivity and Joint Attention. *The International Encyclopedia of Anthropology*, 1-9.

Baldwin, D. A., Markman, E. M., & Melartin, R. L. (1993). Infants' ability to draw inferences about nonobvious object properties: Evidence from exploratory play. *Child development*, *64*(3), 711-728.

Baldwin, D. A., Markman, E. M., Bill, B., Desjardins, R. N., Irwin, J. M., & Tidball, G. (1996). Infants' reliance on a social criterion for establishing word-object relations. *Child development*, *67*(6), 3135-3153.

Baldwin, D.A., Myhr, K., & Brand, R. (in preparation). Motionese elicits higher-fidelity imitation than adult-directed action.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language*, *68*(3), 255-278.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.

Beatty, J., & Lucero-Wagoner, B. (2000). The pupillary system. *Handbook of psychophysiology*, *2*(142-162).

Bettes, B. A. (1988). Maternal depression and motherese: Temporal and intonational features. *Child development*, 1089-1096.

Binda, P., Pereverzeva, M., & Murray, S. O. (2013). Attention to bright surfaces enhances the pupillary light reflex. *Journal of Neuroscience*, *33*(5), 2199-2204.

Brand, R. J., Baldwin, D. A., & Ashburn, L. A. (2002). Evidence for 'motionese': modifications in mothers' infant-directed action. *Developmental Science*, *5*(1), 72-83.

Blount, B. G., & Padgug, E. J. (1977). Prosodic, paralinguistic, and interactional features in parent-child speech: English and Spanish. *Journal of child language*, *4*(1), 67-86.

Brand, R. J., Hollenbeck, E., & Kominsky, J. F. (2013). Mothers' infant-directed gaze during object demonstration highlights action boundaries and goals. *IEEE Transactions on Autonomous Mental Development*, *5*(3), 192-201.

Brand, R. J., McGee, A., Kominsky, J. F., Briggs, K., Gruneisen, A., & Orbach, T. (2009). Repetition in infant-directed action depends on the goal structure of the object: Evidence for statistical regularities. *Gesture*, *9*(3), 337-353.

Brand, R. J., Shallcross, W. L., Sabatos, M. G., & Massie, K. P. (2007). Fine-grained analysis of motionese: Eye gaze, object exchanges, and action units in infant-versus adult-directed action. *Infancy*, *11*(2), 203-214.

Brand, R. J., & Shallcross, W. L. (2008). Infants prefer motionese to adult-directed action. *Developmental science*, *11*(6), 853-861.

Brand, R. J., & Tapscott, S. (2007). Acoustic packaging of action sequences by infants. *Infancy*, *11*(3), 321-332.

Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the society for research in child development*, i-174.

Carter, M. E., Yizhar, O., Chikahisa, S., Nguyen, H., Adamantidis, A., Nishino, S., ... & De Lecea, L. (2010). Tuning arousal with optogenetic modulation of locus coeruleus neurons. *Nature neuroscience*, *13*(12), 1526.

Cooper, R. P., & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child development*, *61*(5), 1584-1595.

Csibra, G., & Gergely, G. (2006). Social learning and social cognition: The case for pedagogy. *Processes of change in brain and cognitive development. Attention and performance XXI*, *21*, 249-274.

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in cognitive sciences*, *13*(4), 148-153.

Csibra, G., & Gergely, G. (2011). Natural pedagogy as evolutionary adaptation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1567), 1149-1157.

Databrary. (2012). The Databrary Project: A video data library for developmental science. New York: New York University.

Datavyu Team. (2014). Datavyu: A Video Coding Tool. Databrary Project, New York University.

Dawson, C., & Gerken, L. (2009). From domain-generality to domain-sensitivity: 4-month-olds learn an abstract repetition rule in music that 7-month-olds do not. *Cognition*, *111*(3), 378-382.

Doupe, A. J., & Kuhl, P. K. (1999). Birdsong and human speech: common themes and mechanisms. *Annual review of neuroscience*, *22*(1), 567-631.

DeBolt, M.C., Rhemtulla, M., & Oakes, L.M. (2019, March 22). Robust data and power in infant looking time research: Number of infants and number of trials. Talk presented at the Society for Research in Child Development Biennial Meeting, Baltimore, MD.

Eschenko, O., & Sara, S. J. (2008). Learning-dependent, transient increase of activity in noradrenergic neurons of locus coeruleus during slow wave sleep in the rat: Brain stem–cortex interplay for memory consolidation?. *Cerebral Cortex*, *18*(11), 2596-2603.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., ... & Stiles, J. (1994). Variability in early communicative development. *Monographs of the society for research in child development*, i-185.

Fenson, L., Marchman, V.A., Thal, D., Dale, P., Reznick, J.S., & Bates, E. (2007). *MacArthur-Bates Communicative Development Inventories: user's guide and technical manual,* 2nd ed. Baltimore, MD: Brookes Publishing Company.

Ferguson, C. A. (1964). Baby talk in six languages. *American anthropologist*, *66*(6), 103-114.

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant behavior and development*, *8*(2), 181-195.

Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of child language*, *16*(3), 477-501.

Flores, S., Bailey, H. R., Eisenberg, M. L., & Zacks, J. M. (2017). Event segmentation improves event memory up to one month later. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(8), 1183.

Foote, S. L., & Morrison, J. H. (1987). Extrathalamic modulation of cortical function. *Annual review of neuroscience*, *10*(1), 67-95.

Frick, J. E., & Richards, J. E. (2001). Individual differences in infants' recognition of briefly presented visual stimuli. *Infancy*, *2*(3), 331-352.

Fukuyama, H., Qin, S., Kanakogi, Y., Nagai, Y., Asada, M., & Myowa-Yamakoshi, M. (2015). Infant's action skill dynamically modulates parental action demonstration in the dyadic interaction. *Developmental science*, *18*(6), 1006-1013.

Geangu, E., Hauf, P., Bhardwaj, R., & Bentz, W. (2011). Infant pupil diameter changes in response to others' positive and negative emotions. *PloS one*, *6*(11), e27132.

Geller, J., Winn, M., Mahr, T., & Mirman, D. (2019). GazeR: A Package for Processing Gaze Position and Pupil Size Data.

Gergely, G., Egyed, K., & Király, I. (2007). On pedagogy. *Developmental science*, *10*(1), 139-146.

Gilzenrat, M. S., Nieuwenhuis, S., Jepma, M., & Cohen, J. D. (2010). Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. *Cognitive, Affective, & Behavioral Neuroscience*, *10*(2), 252-269.

Gleitman, L. R., Newport, E. L., & Gleitman, H. (1984). The current status of the motherese hypothesis. *Journal of child language*, *11*(1), 43-79.

Gogate, L., Maganti, M., & Bahrick, L. E. (2015). Cross-cultural evidence for multimodal motherese: Asian Indian mothers' adaptive use of synchronous words and gestures. *Journal of experimental child psychology*, *129*, 110-126.

Gold, D. A., Zacks, J. M., & Flores, S. (2017). Effects of cues to event segmentation on subsequent memory. *Cognitive research: principles and implications*, *2*(1), 1.

Goldinger, S. D., & Papesh, M. H. (2012). Pupil dilation reflects the creation and retrieval of memories. *Current Directions in Psychological Science*, *21*(2), 90-95.

Goldwater, B.C. (1972). Psychological significance of pupillary movements. *Psychological bulletin, 77(5),* 340.

Gomez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, *13*(5), 431-436.

Gottfried A, Gottfried A, Bathurst K, Wright Guerin D, Parramore M. Socioeconomic status in children's development and family environment: Infancy through adolescence. In: Bornstein M, Bradley R, editors. Socioeconomic status, parenting, and child development. Mahwah: NJ: Lawrence Erlbaum Associates; 2003. pp. 189–207.

Granholm, E., Asarnow, R. F., Sarkin, A. J., & Dykes, K. L. (1996). Pupillary responses index cognitive resource limitations. *Psychophysiology*, *33*(4), 457-461.

Granholm, E., Morris, S. K., Sarkin, A. J., Asarnow, R. F., & Jeste, D. V. (1997). Pupillary responses index overload of working memory resources in schizophrenia. *Journal of Abnormal Psychology*, *106*(3), 458.

Gredebäck, G., & Melinder, A. (2010). Infants' understanding of everyday social interactions: A dual process account. *Cognition*, *114*(2), 197-206.

Hard, B. M., Meyer, M., & Baldwin, D. (2019). Attention reorganizes as structure is detected in dynamic action. *Memory & cognition*, *47*(1), 17-32.

Hard, B. M., Recchia, G., & Tversky, B. (2011). The shape of action. *Journal of experimental psychology: General*, *140*(4), 586.

Hardwicke, T. E., Mathur, M. B., MacDonald, K., Nilsonne, G., Banks, G. C., Kidwell, M. C., ... & Lenne, R. L. (2018). Data availability, reusability, and analytic reproducibility: Evaluating the impact of a mandatory open data policy at the journal Cognition. *Royal Society open science*, *5*(8), 180448.

Hepach, R., Vaish, A., & Tomasello, M. (2012). Young children are intrinsically motivated to see others helped. *Psychological science*, *23*(9), 967-972.

Hepach, R., Vaish, A., & Tomasello, M. (2015). Novel paradigms to measure variability of behavior in early childhood: posture, gaze, and pupil dilation. *Frontiers in psychology*, *6*, 858.

Hepach, R., & Westermann, G. (2013). Infants' sensitivity to the congruence of others' emotions and actions. *Journal of experimental child psychology*, *115*(1), 16-29.

Hepach, R., & Westermann, G. (2016). Pupillometry in infancy research. *Journal of Cognition and Development*, *17*(3), 359-377.

Hespos, S. J., Grossman, S. R., & Saylor, M. M. (2010). Infants' ability to parse continuous actions: Further evidence. *Neural Networks*, *23*(8-9), 1026-1032.

Hespos, S. J., Saylor, M. M., & Grossman, S. R. (2009). Infants' ability to parse continuous actions. *Developmental psychology*, *45*(2), 575.

Hess, E. H., & Polt, J. M. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science*, *143*(3611), 1190-1192.

Hirsh-Pasek, K., & Golinkoff, R. M. (1996). The origins of grammar: Evidence from early language comprehension. Cambridge, MA: MIT Press.

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child development*, *74*(5), 1368-1378.

Hou, R.H., Freeman, C., Langley, R.W., Szabadi, E., & Bradshaw, C.M. (2005). Does modafinil activate the locus coeruleus in man? Comparison of modafinil and clonidine on arousal and automatic functions in human volunteers. *Psychopharmacology, 181*(3), 537-549.

Jackson, I., & Sirois, S. (2009). Infant cognition: going full factorial with pupil dilation. *Developmental science*, *12*(4), 670-679.

Jepma, M., & Nieuwenhuis, S. (2011). Pupil diameter predicts changes in the exploration–exploitation trade-off: Evidence for the adaptive gain theory. *Journal of cognitive neuroscience*, *23*(7), 1587-1596.

Joshi, S., Li, Y., Kalwani, R. M., & Gold, J. I. (2016). Relationships between pupil diameter and neuronal activity in the locus coeruleus, colliculi, and cingulate cortex. *Neuron*, *89*(1), 221-234.

Just, M. A., & Carpenter, P. A. (1993). The intensity dimension of thought: pupillometric indices of sentence processing. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *47*(2), 310.

Kahneman, D. (1973). *Attention and effort* (Vol. 1063). Englewood Cliffs, NJ: Prentice-Hall.

Kahneman, D., & Beatty. J. (1966). Pupil diameter and load on memory. *Science, 154*(3756), 1583-1585.

Kemler-Nelson, D. G., Hirsh-Pasek, K., Jusczyk, P. W., & Cassidy, K. W. (1989). How the prosodic cues in motherese might assist language learning. *Journal of child Language*, *16*(1), 55-68.

Kanakogi, Y., & Itakura, S. (2011). Developmental correspondence between action prediction and motor ability in early infancy. *Nature communications*, *2*, 341.

Kimmerle, M., Mick, L. A., & Michel, G. F. (1995). Bimanual role-differentiated toy play during infancy. *Infant Behavior and Development*, *18*(3), 299-307.

Klein, O., Hardwicke, T. E., Aust, F., Breuer, J., Danielsson, H., Mohr, A. H., ... & Frank, M. C. (2018). A practical guide for transparency in psychological science. *Collabra: Psychology*, *4*(1).

Klimek, V., Stockmeier, C., Overholser, J., Meltzer, H. Y., Kalka, S., Dilley, G., & Ordway, G. A. (1997). Reduced levels of norepinephrine transporters in the locus coeruleus in major depression. *Journal of Neuroscience*, *17*(21), 8451-8458.

Kline, M. A. (2015). How to learn about teaching: An evolutionary framework for the study of teaching behavior in humans and other animals. *Behavioral and Brain sciences*, *38*.

Kline, M. A., Boyd, R., & Henrich, J. (2013). Teaching and the life history of cultural transmission in Fijian villages. *Human Nature*, *24*(4), 351-374.

Kosie, J. E., & Baldwin, D. (2019a). Attention rapidly reorganizes to naturally occurring structure in a novel activity sequence. *Cognition*, *182*, 31-44.

Kosie, J. E., & Baldwin, D. (2019b). Attentional profiles linked to event segmentation are robust to missing information. *Cognitive research: principles and implications*, *4*(1), 8.

Koterba, E. A., & Iverson, J. M. (2009). Investigating motionese: The effect of infant-directed action on infants' attention and object exploration. *Infant Behavior and Development*, *32*(4), 437-444.

Kuhl, P. K. (2003). Human speech and birdsong: communication and the social brain. *Proceedings of the National Academy of Sciences*, *100*(17), 9645-9646.

Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. *Nature reviews neuroscience*, *5*(11), 831.

Kuhl, P. K. (2007). Is speech learning 'gated'by the social brain?. *Developmental science*, *10*(1), 110-120.

Kuhl, P. K., Tsao, F. M., & Liu, H. M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences*, *100*(15), 9096-9101.

Kurby, C. A., & Zacks, J. M. (2008). Segmentation in the perception and memory of events. *Trends in cognitive sciences*, *12*(2), 72-79.

Kurby, C. A., & Zacks, J. M. (2011). Age differences in the perception of hierarchical structure in events. *Memory & cognition*, *39*(1), 75-91.

Kurby, C. A., & Zacks, J. M. (2018). Preserved neural event segmentation in healthy older adults. *Psychology and aging*, *33*(2), 232.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13).

Laeng, B., Sirois, S., & Gredebäck, G. (2012). Pupillometry: A window to the preconscious?. *Perspectives on psychological science*, *7*(1), 18-27.

Lenth, R.V. (2016). Least-Squares Means: The R Package lsmeans. Journal of Statistical Software, 69(1), 1-33.

Levine, D., Buchsbaum, D., Hirsh-Pasek, K., & Golinkoff, R. M. (2019). Finding events in a continuous world: A developmental account. *Developmental psychobiology*, *61*(3), 376-389.

Liaw, F. R., & Brooks-Gunn, J. (1994). Cumulative familial risks and low-birthweight children's cognitive and behavioral development. *Journal of Clinical Child Psychology*, *23*(4), 360-272.

Lockman, J. J., & McHale, J. P. (1989). Object manipulation in infancy. In *Action in Social Context* (pp. 129-167). Springer, Boston, MA.

Loucks, J., & Baldwin, D. (2009). Sources of information for discriminating dynamic human actions. *Cognition*, *111*(1), 84-97.

Loewenfeld, E. (1993). The pupil: Anatomy, physiology, and clinical applications. Detroit: Wayne State University Press.

Luke, S. G. (2017). Evaluating significance in linear mixed-effects models in R. *Behavior research methods*, *49*(4), 1494-1502.

ManyBabies Consortium (under revision). Quantifying sources of variability in infancy research using the infant-directed speech preference. *Advances in Methods and Practices in Psychological Science.*

Mathôt, S., Fabius, J., Van Heusden, E., & Van der Stigchel, S. (2018). Safe and sensible preprocessing and baseline correction of pupil-size data. *Behavior research methods*, *50*(1), 94-106.

MATLAB and Statistics Toolbox Release 2019a, The MathWorks, Inc., Natick, Massachusetts, United States.

Martineau, J., Hernandez, N., Hiebel, L., Roché, L., Metzger, A., & Bonnet-Brilhault, F. (2011). Can pupil size and pupil responses during visual scanning contribute to the diagnosis of autism spectrum disorder in children?. *Journal of psychiatric research*, *45*(8), 1077-1082.

Meyer, M., Hard, B., Brand, R. J., McGarvey, M., & Baldwin, D. A. (2011). Acoustic packaging: Maternal speech and action synchrony. *IEEE Transactions on Autonomous Mental Development*, *3*(2), 154-162.

Miller, A. L., Gross, M. P., & Unsworth, N. (2019). Individual differences in working memory capacity and long-term memory: The influence of intensity of attention to items at encoding as measured by pupil dilation. *Journal of Memory and Language*, *104*, 25-42.

Monroy, C., Gerson, S., & Hunnius, S. (2017). Infants' motor proficiency and statistical learning for actions. *Frontiers in psychology*, *8*, 2174.

Morad, Y., Lemberg, H., Yofe, N., & Dagan, Y. (2000). Pupillography as an objective indicator of fatigue. *Current eye research*, *21*(1), 535-542.

Morita, T., Slaughter, V., Katayama, N., Kitazaki, M., Kakigi, R., & Itakura, S. (2012). Infant and adult perceptions of possible and impossible body movements: An eye-tracking study. *Journal of Experimental Child Psychology*, *113*(3), 401-414.

Murphy, P. R., Robertson, I. H., Balsters, J. H., & O'connell, R. G. (2011). Pupillometry and P3 index the locus coeruleus–noradrenergic arousal function in humans. *Psychophysiology*, *48*(11), 1532-1543.

Nassar, M. R., Rumsey, K. M., Wilson, R. C., Parikh, K., Heasly, B., & Gold, J. I. (2012). Rational regulation of learning dynamics by pupil-linked arousal systems. *Nature neuroscience*, *15*(7), 1040.

Neville, H., Stevens, C., Pakulak, E., & Bell, T. A. (2013). Commentary: Neurocognitive consequences of socioeconomic disparities. *Developmental science*, *16*(5), 708-712.

Newtson, D. (1973). Attribution and the unit of perception of ongoing behavior. *Journal of Personality and Social Psychology*, *28*(1), 28.

Nieuwenhuis, S., De Geus, E. J., & Aston-Jones, G. (2011). The anatomical and functional relationship between the P3 and autonomic components of the orienting response. *Psychophysiology*, *48*(2), 162-175.

Noble, K. G., McCandliss, B. D., & Farah, M. J. (2007). Socioeconomic gradients predict individual differences in neurocognitive abilities. *Developmental science*, *10*(4), 464-480.

Nuske, H. J., Vivanti, G., Hudry, K., & Dissanayake, C. (2014). Pupillometry reveals reduced unconscious emotional reactivity in autism. *Biological psychology*, *101*, 24-35.

Nuske, H. J., Vivanti, G., & Dissanayake, C. (2015). No evidence of emotional dysregulation or aversion to mutual gaze in preschoolers with autism spectrum disorder: an eye-tracking pupillometry study. *Journal of autism and developmental disorders*, *45*(11), 3433-3445.

Oakes, L. M. (2017). Sample size, statistical power, and false conclusions in infant looking-time research. *Infancy*, *22*(4), 436-469.

Peavler, W. S. (1974). Pupil size, information overload, and performance differences. *Psychophysiology*, *11*(5), 559-566.

Peirce, J. W. (2007). PsychoPy—psychophysics software in Python. Journal of Neuroscience Methods, 162(1), 8–13.

Poynton, C.A., (2003). Digital video and HDTV: Algorithms and interfaces. San Francisco, CA: Morgan Kaufmann.

Preuschoff, K., t Hart, B. M., & Einhauser, W. (2011). Pupil dilation signals surprise: Evidence for noradrenaline's role in decision making. *Frontiers in neuroscience*, *5*, 115.

Putnam, S. P., Helbig, A. L., Gartstein, M. A., Rothbart, M. K., & Leerkes, E. (2014). Development and assessment of short and very short forms of the Infant Behavior Questionnaire–Revised. *Journal of personality assessment*, *96*(4), 445-458.

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Radvansky, G. A., & Zacks, J. M. (2017). Event boundaries in memory and cognition. *Current opinion in behavioral sciences*, *17*, 133-140.

Rajkowski, J., Majczynski, H., Clayton, E., & Aston-Jones, G. (2004). Activation of monkey locus coeruleus neurons varies with difficulty and performance in a target detection task. *Journal of Neurophysiology*, *92*(1), 361-371.

Richards, J. E. (1997). Effects of attention on infants' preference for briefly exposed visual stimuli in the paired-comparison recognition-memory paradigm. *Developmental Psychology*, *33*(1), 22.

Richmond, L. L., Gold, D. A., & Zacks, J. M. (2017). Event perception: Translations and applications. *Journal of Applied Research in Memory and Cognition*, *6*(2), 111-120.

Rochat, P., Passos-Ferreira, C., & Salem, P. (2009). Three levels of intersubjectivity in early development. *Enacting intersubjectivity: Paving the way for a dialogue between cognitive science, social cognition and neuroscience*, 173-190.

Rohlfing, K. J., Fritsch, J., Wrede, B., & Jungmann, T. (2006). How can multimodal cues from child-directed interaction reduce learning complexity in robots?. *Advanced Robotics*, *20*(10), 1183-1199.

Roseberry, S., Richie, R., Hirsh-Pasek, K., Golinkoff, R. M., & Shipley, T. F. (2011). Babies catch a break: 7-to 9-month-olds track statistical probabilities in continuous dynamic events. *Psychological Science*, *22*(11), 1422-1424.

Rothbart, M. K. (1981). Measurement of temperament in infancy. *Child development*, 569-578.

Roy, B. C., Frank, M. C., DeCamp, P., Miller, M., & Roy, D. (2015). Predicting the birth of a spoken word. *Proceedings of the National Academy of Sciences*, *112*(41), 12663-12668.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926-1928.

Sage, K. D., & Baldwin, D. (2010). Social gating and pedagogy: Mechanisms for learning and implications for robotics. *Neural Networks*, *23*(8-9), 1091-1098.

Sara, S.J. (2009). The locus coeruleus and noradrenergic modulation of cognition. *Nature reviews neuroscience, 10*(3), 211.

Sargent, J. Q., Zacks, J. M., Hambrick, D. Z., Zacks, R. T., Kurby, C. A., Bailey, H. R., ... & Beck, T. M. (2013). Event segmentation ability uniquely predicts event memory. *Cognition*, *129*(2), 241-255.

Saylor, M. M., Baldwin, D. A., Baird, J. A., & LaBounty, J. (2007). Infants' on-line segmentation of dynamic human action. *Journal of Cognition and Development*, *8*(1), 113-128.

Schieffelin, B. B. (1979). Getting it together: An ethnographic approach to the study of the development of communicative competence. *Developmental pragmatics*, 73-108.

Schluroff, M. (1983). In the eye of the beholder: Cognitive effort during sentence processing. *Meaning, use, and interpretation of language*, 302-323.

Silberzahn, R., Uhlmann, E. L., Martin, D. P., Anselmi, P., Aust, F., Awtrey, E., ... & Carlsson, R. (2018). Many analysts, one data set: Making transparent how variations in analytic choices affect results. *Advances in Methods and Practices in Psychological Science*, *1*(3), 337-356.

Singmann, H., Bolker, B., Westfall, J., & Aust, F. (2017). afex: Analysis of factorial experiments. R package version 0.21-2.

Sirois, S., & Brisson, J. (2014). Pupillometry. *Wiley Interdisciplinary Reviews: Cognitive Science*, *5*(6), 679-692.

Sirois, S., & Jackson, I.R. (2012). Pupil dilation and object permanence in infants. *Infancy, 17*(1), 61-78.

Snow, C. E., & Ferguson, C. A. (1977). Talking to children.

Sokolov, E. N. (1963). Higher nervous functions: The orienting reflex. *Annual review of physiology*, *25*(1), 545-580.

Sommerville, J. A., Woodward, A. L., & Needham, A. (2005). Action experience alters 3-month-old infants' perception of others' actions. *Cognition*, *96*(1), B1-B11.

Sonne, T., Kingo, O. S., & Krøjgaard, P. (2017). Bound to remember: Infants show superior memory for objects presented at event boundaries. *Scandinavian journal of psychology*, *58*(2), 107-113.

Sonne, T., Kingo, O. S., & Krøjgaard, P. (2016). Occlusions at event boundaries during encoding have a negative effect on infant memory. *Consciousness and cognition*, *41*, 72-82.

Stahl, A. E., Romberg, A. R., Roseberry, S., Golinkoff, R. M., & Hirsh-Pasek, K. (2014). Infants segment continuous events using transitional probabilities. *Child development*, *85*(5), 1821-1826.

Sterpenich, V., D'Argembeau, A., Desseilles, M., Balteau, E., Albouy, G., Vandewalle, G., ... & Maquet, P. (2006). The locus ceruleus is involved in the successful retrieval of emotional memories in humans. *Journal of Neuroscience*, *26*(28), 7416-7423.

Swallow, K. M., Zacks, J. M., & Abrams, R. A. (2009). Event boundaries in perception affect memory encoding and updating. *Journal of Experimental Psychology: General*, *138*(2), 236.

Tanaka, Y., Kosie, J. E., & Baldwin, D. (in prep). Implicit measure of event segmentation using pupillary response.

Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy*, *7*(1), 53-71.

Trainor, L. J., Austin, C. M., & Desjardins, R. N. (2000). Is infant-directed speech prosody a result of the vocal expression of emotion?. *Psychological science*, *11*(3), 188-195.

Trevarthen, C. (1977). Descriptive analyses of infant communicative behaviour. *Studies in mother-infant interaction*.

Trevarthen, C., & Hubley, P. (1978). Secondary intersubjectivity: Confidence, confiding and acts of meaning in the first year. Action, gesture and symbol. The emergence of language. *A. Lock. New York: Academic*, 183-229.

Unsworth, N., & Robison, M. K. (2015). Individual differences in the allocation of attention to items in working memory: Evidence from pupillometry. *Psychonomic Bulletin & Review*, *22*(3), 757-765.

Verschoor, S. A., Spapé, M., Biro, S., & Hommel, B. (2013). From outcome prediction to action selection: developmental change in the role of action–effect bindings. *Developmental Science*, *16*(6), 801-814.

Verschoor, S. A., Paulus, M., Spapé, M., Biro, S., & Hommel, B. (2015). The developing cognitive substrate of sequential action control in 9-to 12-month-olds: Evidence for concurrent activation models. *Cognition*, *138*, 64-78.

Weiskrantz, L., Cowey, A., & Le Mare, C. (1998). Learning from the pupil: a spatial visual channel in the absence of V1 in monkey and human. *Brain: a journal of neurology*, *121*(6), 1065-1072.

Weiskrantz, L., Cowey, A., & Barbur, J. L. (1999). Differential pupillary constriction and awareness in the absence of striate cortex. *Brain*, *122*(8), 1533-1538.

Werker, J. F., & McLeod, P. J. (1989). Infant preference for both male and female infant-directed talk: a developmental study of attentional and affective responsiveness. *Canadian Journal of Psychology/Revue canadienne de psychologie*, *43*(2), 230.

Williamson, R. A., & Brand, R. J. (2014). Child-directed action promotes 2-year-olds' imitation. *Journal of experimental child psychology*, *118*, 119-126.

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, *69*(1), 1-34.

Zacks, J. M., Speer, N. K., Vettel, J. M., & Jacoby, L. L. (2006). Event understanding and memory in healthy aging and dementia of the Alzheimer type. *Psychology and aging*, *21*(3), 466.

Zacks, J. M., Tversky, B., & Iyer, G. (2001). Perceiving, remembering, and communicating structure in events. *Journal of experimental psychology: General*, *130*(1), 29.

Zalla, T., Labruyère, N., & Georgieff, N. (2013). Perceiving goals and actions in individuals with autism spectrum disorders. *Journal of autism and developmental disorders*, *43*(10), 2353-2365.

Zhang, F., Jaffe-Dax, S., Wilson, R. C., & Emberson, L. L. (2019). Prediction in infants and adults: A pupillometry study. *Developmental science*, *22*(4), e12780.